Citation: Kalogeropoulos, G. E. (1985). Matrix pencils and linear system theory. (Unpublished Doctoral thesis, The City University)

THE CITY UNIVERSITY
LONDON EC1VOHB
SCHOOL OF ELECTRICAL ENGINEERING
AND APPLIED PHYSICS
CONTROL ENGINEERING CENTRE

Thesis Submitted
for the Award of the Degree
of PH.D in Mathematical
Control Theory

MATRIX PENCILS AND LINEAR SYSTEM THEORY

by

Grigoris El. Kalogeropoulos

October 1985

Στη γυναίκα μου Γεωργία
και στα παιδιά μου Ηλιάνα και Βάσ

DECLARATION

The University Librarian of the City University may allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies, made for study purposes, subject to normal conditions of acknowledgement.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincerest gratitude to [redacted]

I would also like to express my thanks to my friend and research colleague Dr Theodore [illegible] for many stimulating discussions and comments related to this research.

I am also grateful to the Mathematics Department of the University of Athens for granting me two-years leave and for their understanding. Especially, I would like to thank Professor A. [illegible] for his continuous support and interest in the progress of my work.

Thanks are also due to Mrs P. [illegible] who has excelled her in typing the difficult text of this thesis.

I wish to express my gratitude to my mother [illegible]; I thank her not only for her warm hospitality, but also for her general support and understanding during this difficult moments of the period of my studies.

My postgraduate studies, leading to this thesis, have been made possible by the unbounded support of my wife, [illegible]. Without her undertaking the full share of responsibilities for our family, for a period of three years, my studies abroad could never materialise. In recognition of her invaluable contributions, I dedicate this thesis to her.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincerest gratitude to my thesis suprevisor and systems theory teacher Dr Nicos Karcanias. With a wealth of prodding questions, assignements, suggestions, and discussions, with generous help, support, and encouragement, with shred and attentive guidance in short, he has worked at least as hard as I have to bring out the best in me. His enthousiasm, experience and active interest in this research area have helped me survive those periods of despair that few doctoral students escape. For all this, and more, I shall always be grateful.

I would also like to express my thanks to my friend and research coleague [redacted] for many stimulating discussions and comments related to this research.

I am very grateful to the Mathematics Department of the University of Athens for granting me the study leave and for their understanding. Especially, I would like to thank [redacted] for his continuous support and interest in the progress of my work.

Thanks are also due to [redacted] for her excellent jop in typing the largest part of this thesis.

I wish to express my grattitude to my sister [redacted]; I thank her, not only for her warm hospitality, but also for her general support and understanding during many difficult moments in the period of my studies.

My postgraduate studies, leading to this thesis, have been made possible by the unbounded support of my wife, [redacted]. Without her undertaking the full range of responsibilities for our family, for a period of three years, my studies abroad could never matterialise. In recognition of her invaluable contributions, I dedicate this thesis to her.

ABSTRACT

The study of problems and structural properties of regular and extended state space theory may be reduced to a study of first-order linear differential equations

$$S(F,G): F\underline{\dot{x}}(t) = G\underline{x}(t), \quad F, G \in \mathbb{R}^{mxn}$$

Matrix pencil theory is the key tool for the study of $S(F,G)$ differential systems. The developement of a general theory for $S(F,G)$ systems, and thus the developement of a unifying theory for linear systems, necessitates the further enrichment of the classical theory of strict equivalenve for matrix pencils. To cover the needs of linear systems, a matrix pencil theory has to be general enough and should have a geometric, dynamic, topological, invariant theory and computational dimension. This thesis aspires to contribute in the developement of the matrix pencil theory along the above lines and thus contribute in the foundations of a matrix pencil based unifying theory for linear systems.

The theory of strict equivalence is detached from its algebraic context and it is presented as a theory of ordered pairs (F,G). The strict equivalence invariants are defined in a number theoretic way, by the properties of appropriate Peicewise Arithmetic progression sequences defined on a pair (F,G). This new characterisation of strict equivalence invariants allows the derivation of new procedures for computing the Kronecker canonical form, for constructing minimal bases, and provides the means for the establishment of the geometric theory of strict equivalence invariants. The subspaces of the domain of (F,G) are classified in terms of the invariants of the restriction pencil. A variety of notions of invariant subspaces emerges, such as (F,G)- , (G,F)-, complete-(F,G)-invariant subspaces and extended-(F,G)-, (G,F)-, complete-(F,G)-invariant subspaces. These notions of invariant subspaces are

the counterparts of the standard notions of invariant subspaces of the geometric theory. The properties of the solution space of $S(F,G)$ differential systems are studied and the different norions of invariant subspaces are characterised dynamically in terms of the properties of C^{∞}-, distributional-holdability and C^{∞}-, distributional reachability. Once more, these dynamic properties are generalisations of the fundamental notions of geometric theory. The theory of invatiants of matrix pencils, or ordered pairs, is enriched by the study of invariants under Bilinear strict equivalence; a complete set of invariants is defined under this equivalence. This study leads to a "space frequency" relativistic classification of the dynamic and geometric properties of $S(F,G)$ systems and their invariant subspaces; furthermore, it provides the means for a systematic study of dual systems and problems in linear systems. The further developement of the theory of invariant forced realisations, allows the translation of results and properties derived on $S(F,G)$ back to linear systems theory. Finally, the problem of defining appropriate topological settings for the study of properties of pencils under under uncertainty in their description is examined. New metric topologies are introduced and their links to known results of the perturbation theory of the generalised eigenvalue-eigenvector problem are established.

NOTATION AND ABBREVIATIONS

Throughout this thesis, the following notation and abbreviations will be used:

$\mathbb{R},\mathbb{C},\mathbb{R}[s]$	the field of real, complex numbers and rational functions, respectively
$\mathbb{Z},\mathbb{N}$	the set of integers and natural numbers, respectively
$\mathbb{R}[s]$	the ring of polynomials over $\mathbb{R}$
$\mathbb{R}^n,\mathbb{C}^n,\mathbb{R}^n(s)$	the n-dimensional vector spaces over $\mathbb{R},\mathbb{C},\mathbb{R}(s)$
$\mathbb{F}^{mxn}$	the set of mxn matrices with elements from the field $\mathbb{F}$
$PGL(1,\mathbb{C})$	the general projective group on the projective straight line of $\mathbb{C}\cup\{\infty\}$
$D(0),D(\infty),D(a)$	the set of zero, infinite and a elementary divisors, respectively
$I_c(F,G)$	the set of column minimal indices of $sF-\hat{s}G$
$I_r(F,G)$	the set of row minimal indices of $sF-\hat{s}G$
E_H	the strict equivalence
E_{H-B}	the bilinear-strict equivalence
$E_H(F,G)$	the strict equivalence class of $L=(F,G)$
$P^k_{a}{}^{*}(F,G)$	the a-Toeplitz matrices which are specified by a and the pair (F,G)
N^k_a	denote a nested basis matrix of $N^k_a=N_r\{P^k_a(F,G)\}$
$S_a(F,G),W_a(F,G)$	the Segre, Weyr characteristics of $sF-G$ at $s=a$, respectively
$E_p(f)$	the projective equivalence class of $f(s,\hat{s})$
$i\in\underline{r}$	$i\in\{1,2,\ldots,r\}$
$[B^\pi_{\mathbb{C}}(f)],[B^\pi_{\mathbb{R}}(f)]$	the $(\mathbb{C}-\pi)$-$(\mathbb{R}-\pi)$ basis matrix of $f(s,\hat{s})$, respectively
$J_{\mathbb{C}}(f)$	the complex list of $f(s,\hat{s})$
$J_{\mathbb{R}}(f)$	the real list of $f(s,\hat{s})$

$[\mathcal{B}^{\pi,\pi'}(f)]$	$(\pi-\pi')$-basis matrix of $f(s,\hat{s})$
$\mathcal{E}^{c}_{eh},\mathcal{E}^{r}_{eh}$	the complex-real Hermite equivalence, respectively
$\Lambda^{p}\mathcal{V}$	the p-th exterior power of the vector space $\mathcal{V}$
$Q_{m,n}$	the set of lexicographically ordered, strictly increasing sequences of m integers from 1,2,...,n
$\underline{a}_{i_1}\wedge\dots\wedge\underline{a}_{i_m}=\underline{a}_{\omega}\wedge$	the exterior product of the vectors $\underline{a}_1,\dots,\underline{a}_m$ of an n-dimensional vector space $\mathcal{V}$ where $\omega=(i_1,\dots,i_m)\in Q_{m,n}$
$C_p(A)$	the p-th compound matrix of $A\in F^{m\times n}$, $p\leq\min(m,n)$
ϕ_r	the set of quadruples
P_r	the set of r-prime Plücker vectors of T
$L(\mathcal{V};W)$	the set of all linear mappings from $\mathcal{V}$ into W
c.-(F,G)-i.s.	complete-(F,G)-invariant subspace
(F,G)-i.s.	(F,G)-invariant subspace
M_a^i	i-th generalised null space of (F,G) at s=a
M_a^*	maximal generalised null space of (F,G) at s=a
$S_a(F,G)$	the Segre characteristic of (F,G) at s=a
$\mathcal{W}_a(F,G)$	the Weyr characteristic of (F,G) at s=a
$\tilde{\Sigma}_a(F,G)$	normal complete prime set of chains of (F,G) at s=a
$L_{\varepsilon}(s,\hat{s})$	the standard c.m.i. block associated with ε
$\mathcal{W}_r$	right Weyr sequence of (F,G)
$P_k[(s,\hat{s});N_k]$	N_k-right annihilating set of $sF-\hat{s}G$
$M_{\sigma_1},\hat{M}_{\sigma_1}$	$\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$-, σ_1-right annihilating modules of (F,G) respectively
$\mathcal{W}^{i}(N_k)$	ith supporting space of N_k
$K_r(F,G)$	right set of singularity of (F,G)
$L_{m,n}$	the set: $\{(F,G): F,G\in\mathbb{R}^{m\times n}\}$
APR	arithmetic progression relationship
APS	arithmetic progression sequence

BE	bilinear equivalence
BSEG	bilinear strict equivalence group
BSE	bilinear strict equivalence
BEG	bilinear equivalence group
(ℂ,UFS)	complex unique factorisation set
c.i.v.	consistent initial vector
d.g.s.	dual generalised state
d.s.t.	distributional state trajectory
e.d.	elementary divisor
e.r.o.'s	elementary row operations
e.c.o.'s	elementary column operations
e.l.s.	entirely left singular
e.r.s.	entirely right singular
e.e.r.s.	extended entirely right singular
e.e.r.r.	extended entirely right regular
f.e.d.	finite elementary divisors
g.c.d.	greatest common divisor
G.R.D.	greatest right divisor
G.L.D.	greatest left divisor
G.C.R.D.	greatest common right divisor
G.C.L.D.	greatest common left divisor
G.L.M.P.	general linear mapping problem
g.s.	generalised state
(K-a-(F,G)_T.M.)	K-th order a-(F,G)-Toeplitz matrices
L.D.	left divisor
M.F.D.	matrix fraction description
m.r.f.i.r.	minimal regular forced invariant realisation
m.r.i.r.	minimal regular invariant realisation

(NI.PAPS)	non-increasing PAPS
n.K.o.	natural Krönecker orbit
n.B.o.	natural Brunovsky orbit
n.s.t.	normal state trajectory
O.I.M.	orthogonal invariant metric
PAPS	piecewise arithmetic progression sequence
PAPSD	piecewise arithmetic progression diagram
p.r.f.i.r.	proper regular invariant realisations
RPAPS	right piecewise arithmetic progression sequence
($\mathbb{R}$-UFS)	real unique factorisation set
r.f.i.r.	regular forced invariant realisation
r.i.r.	regular invariant realisation
SEG	strict equivalence group
TVS	topological vector space
UFS	unique factorisation set
WSD	Weyr sequence diagram
z.e.d.	zero elementary divisors

CONTENTS

CHAPTER 1:

Introduction

CHAPTER 1: INTRODUCTION

The theory that will be presented in this thesis falls within an area of research that is referred to as "matrix pencil approach" to the geometric theory of linear systems. The two fundamental parts of this approach are the geométric theory of linear systems and the algebraic theory of matrix pencils. The work presented here is an attempt to unify the geometric, algebraic and dynamic aspects of linear system problems, which may be reduced to the study of properties of generalised autonomous differential systems $S(F,G):F\dot{\underline{x}}=G\underline{x}$, $F, G \in \mathbb{R}^{mxn}$.

The basic philosophy underlying the geometric approach to linear systems, is that a system is an entity defined by a number of mappings defined on abstract linear spaces (the input, the state and the output space); several relevant structural features of the system are therefore determined by the way in which these mappings interwine in their domains and codomains. These structural features can be expressed in terms of the geometrical properties of different types of subspaces connected with these mappings. Questions involving the existence and synthesis of controllers may be reduced to problems concerning the interelationships of certain subspaces and the existence of mappings with given properties. Of key importance to this approach is the characterisation and properties of various different types of invariant subspaces. The basic references for the geometric approach may be found in [Won. -1],[Will. -1].

The matrix pencil approach to geometric theory [Was. & Eck. -1],[MacF. & Kar. -1],[Kar. -1],[Jaf. & Kar. -1] has been motivated by the need to simplify the characterisation of the basic concepts and the solution of problems of the geometric theory, by introducing an algebraic dimension to them. The characterisation of subspaces of the state space is based on the strict equivalence invariants of a special pencil, called the "restriction pencil" of the subspace. This approach has provided the means for a better understanding of the structural aspects of linear systems and it has reduced the problem of computations to the study of a generalised eigenvalue-eigenvector problem. The present research is motivated by the results of the standard matrix pencil approach. It aspires to contribute in the developement of a matrix pencil approach for a broader class of linear systems, that includes both regular and extended state space systems. Of crucial importance for such an attempt is the developement of a geometric and dynamic theory of matrix pencils, or of autonomous generalised differential systems. The main emphasis of the present work is on the developement of a general theory of $S(F,G)$ systems.

The geometric approach for linear regular state space systems is a theory developed for first order non-singular differential equations. The intimate link of general first order linear differential equations with matrix pencil operators, suggests that the matrix pencil theory is the natural setting for the developement of the geometric, algebraic, dynamic and computational aspects of broad classes of linear dynamical systems; within the families of linear systems that may be studied in terms of matrix pencils, we distinguish the singular, or extended state space systems, regular state space systems and systems with uncertainty in their parameters. In the attempt to generalise the standard geometric theory to singular systems we face a number of difficulties; these are due to the lack of a geometric theory of matrix pencils and to the obstacles in the interpretation of strict equivalence transformations as meaningfull feedback operations on the system. Extending the standard theory to linear systems with uncertainty in the description of the basic maps, becomes a more difficult problem; apart from results on the perturbation theory of eigenvalues, there is no general theory of topological properties of the S.E. invariants and associated subspaces for matrix pencils.

A special case of "uncertain" descriptions is the family of singularly perturbed systems; for this case, the natural operator becomes the "matrix combinant", or "matrix net", $\sum_{i=1}^{k} \lambda_i F_i$, $F_i \in \mathbb{R}^{mxn}$, λ_i indeterminates. A theory for such operators has not been developed yet. The algebraic theory of S.E. of matrix pencils is also inadequate in dealing with problems involving "space-frequency" transformations on matrix pencils and which arise in the study of dual systems and problems [Kar. & Hay. -1,3], in the "space-frequency relativistic" classification of system properties, as well as in the conditioning of the generalised eigenvalue-eigenvector problem [Kubl. -1].

It is evident, that a prerequisite of the developement of a general theory for linear systems is the further enrichment of matrix pencil theory beyond the classifical theory of S.E. and the recent numerical analysis advances [Kag. & Ruh. -1]. The aim of the present work is to contribute in the developement of those aspects of matrix pencil theory, which primarily concern linear systems theory. The main topis discussed are:

(1) Number theoretic aspects of S.E. invariants.

(2) Geometric theory of matrix pencils.

(3) Bilinear-Strict equivalence of matrix pencils.

(4) Dynamic aspects of $S(F,G)$ differential systems.

(5) Framework for the study of topological properties of matrix pencils.

These topics are of key importance for the developement of a general theory of generalised autonomous differential systems that encompasses the algebraic, geometric, dynamic, "relativistic", topological and compututional aspects. A general theory for such differential systems, may provide the means for a unifying and general approach for linear systems of the regular and extended state space type.

The standard eigenvalue-eigenvactor problem defined on a matrix $A \in \mathbb{R}^{nxn}$ has algebraic, geometric and number theoretic aspects. The algebraic properties are defined by the S.E. invariants of the pencil sI-A, the geometric properties are connected with the structural aspects of generalised null spaces (elementary A-invariant subspaces) and the number theoretic properties are implicit in the Segrè, Weyr characteristic theory [Tur. & Ait. -1]. Note that each one of the above properties, and quite independently from the others, may be used for the computation of the Jordan form and thus for defining the structure of the eigenvalue-eigenvector problem. This implies that the theory of the eigenvalue-eigenvector problem may be presented on the pair (I,A) without using the algebraic structure of the pencil sI-A.

Generalising the geometric and number theoretic properties from sI-A to the general sF-G case, implies also a detachment of the theory of S.E. invariants from its purely algebraic basis (Smith form, theory of minimal bases [For. -1]); this is equivalent to presenting the theory on the ordered pairs (F,G) without resorting to the algebraic notion of the matrix pencil. A theory of S.E. presented on (F,G) has a number of important advantages from the conceptual and computational points of view. In fact, similarly to the (I,A) simple case, the number theoretic properties, expressed by an extended Segrè-Weyr theory, may be used for the definition of the set of S.E. invariants of (F,G) and thus for the computation of the Kronecker form. The computational advantages, within the framework of the new definitions, stem from that standard numerical linear algebra techniques may be deployed. Extending the geometric properties from (I,A) to the general (F,G) case, it is expected that a variety of new notions of invariant subspaces, rich in dynamic properties will emerge; the standard matrix pencil approach and geometric theory indicate the existence of such subspaces, which are rich in dynamic properties and thus may describe a number of important properties in linear systems. Finally, we should point out, that a S.E. theory presented on pairs (F,G) provides a more convenient setting for generalisations of the theory, when it is compared with the algebraic approach based on sF-G.

Frequency domain transformations of the bilinear type, have been used in the study of various problems of linear systems theory, as well as in numerical analysis. In the context of linear systems, special bilinear transformations have been deployed for the definition of infinite zeros of a rational matrix [Verg. -1], for the computation of finite and infinite zeros of a rational matrix [Kouv. & Edm. -1] and the definition and study of properties of the "integrator-differentiator" type of duality [Kar. & Hay. -1,3]. In the context of numerical analysis, special bilinear transformations have been used for improving the bad-conditioning of the generalised eigenvalue-eigenvector problem [Kubl. -1]. The need for a general theory "space-frequency" domain transformations is apparent.

The origins of Bilinear-Strict Equivalence (B.S.E.) go back to the classical theory of matrix pencils [Tur. & Ait. -1]. However, apart from some preliminary results, there is no general theory of invariants under B.S.E. Given that B.S.E. expresses coordinate transformations in the space (domain and codomain of a pair (F,G)) and frequency (coordinate transformations on the Riemann sphere) domains, such a theory has a "relativistic", space-frequency dimension. From the theoretical viewpoint, a general theory of invariants under B.S.E. is a prerequisite for the developement of a theory of dual problems, as well as for the "relativistic" classification of important linear systems concepts. Thus, such a theory may provide the means for a "relativistic" classification of system theoretic notions, such as stability, controllability, observability, (A,B)-invariance e.t.c., according to their properties to vary, or remain invariant under "space-frequency" transformations. From the numerical analysis viewpoint, it is believed that such a theory is necessary for the "optimal" conditioning of the generalised eigenvalue-eigenvector problem. The basic idea behind the latter problem, is the definition of a suitable equivalent problem, with "nice" computational properties; the computations may then be carried out on the equivalent problem and the results are then interpreted back to the original badly conditioned problem.

The importance of the study of dynamic aspects of the autonomous genralised differential system $S(F,G)$, stems from its links to the study of motions, which are restricted in a given subspace of the state space of a linear regular, or singular system. This study, provides the means for a dynamic characterisation of the various types of invariant subspaces of the domain of (F,G); it is therefore, instrumental in generalising the fundamental notions of (A,B)-invariant, almost controllability subspaces of linear systems, to the case of $S(F,G)$ differential systems. By establishing an expression for these fundamental dynamical concepts, on the more abstract level of $S(F,G)$ descriptions, the problem of their translation to particular cases, such extended state space systems, becomes much easier.

There are certain conceptual difficulties associated with the $S(F,G)$ differential systems. In fact, in the general case, there exist certain initial conditions for which the uniqueness property of the solution does not hold true; thus, $S(F,G)$ descriptions do not always represent a dynamical system. To overcome these conceptual obstacles, a theory is needed to parametrise the solutions and explain the arbitrariness in a meaningfull system theoretic way. The theory of "invariant forced realisations" of the $S(F,G)$ differential systems [Kar. & Hay. -2], has been introduced to serve that purpose. In fact, it allows the parametrisation of trajectories in a family of a given initial condition, in terms of external control inputs; furthermore, it allows the interpretation of $S(F,G)$ as a "feedback free" description of a control problem defined on an orbit of linear systems.

The necessity for an appropriate topological framework for studying the properties of a pair (F,G), or the pencil sF-G needs hardly to be emphasised. The uncertainty in the exact values of parameters of real life linear models is inherent; two of the main reasons are inaccuracies in the experimental data and rounding off errors in computations. Implicit on the study of robustness of the algebraic and geometric properties of a pair (F,G), is the setting up of an appropriate topological framework; for a topological framework to be suitable for the study of robustness it must have the following properties: It must be natural, as far as describing the origins of uncertainty, should be related to the fundamental properties of (F,G), and must be strong enough for the derivation of strong robustness results.

The origins of the theory presented in this thesis, come from the work of Karcanias [Kar. & Hay. -1,2],[Kar. -5] on the theory of generalised autonomous differential systems $S(F,G)$. This thesis aspires to contribute in the developement of the proper foundations of a general theory of $S(F,G)$ type differential systems; such a theory is considered as a prerequisite of a matric pencil approach for broader families of linear systems. The material presented here is structured in the following way.

The second chapter of this thesis provides a selective summary of concepts results, and tools from the mathematical topics which are essential for the developement of the results in the following chapters. The range of topics discussed there are from: the theory of invariants, polynomial matrices and rational vector spaces theory, the classical theory of matrix pencils, the Segrè-Weyr characteristic theory of the eigenvalue-eigenvector problem, Exterior algebra and Compound matrix theory, and finally, the essential topological notions.

The third chapter "set the scene" for the theory that will be developed in the following chapters. Section (3.2) provides a brief summary of the fundamental concepts of geometric theory, as they have been developed [Won. -1],[Will. -1]. The emphasis is on the dynamic and geometric characterisation of (A,B)-invariant, almost (A,B)-invariant, controllability, almost controllability subspaces. In the same section, a brief summary of the matrix pencil approach to geometric theory [Kar. -1],[Jaf & Kar. -1] is also given; the emphasis there is on the algebraic characterisation of the fundamental invariant subspaces. Section (3.3) shows how matrix pencil theory and the generalised autonomous differential systems $\dot{S}(F,G)$, arise in the context of regular and extended state space systems. It is there, where the $S(F,G)$ description emerges as the unifying object for the study of structural properties of linear systems. In section (3.4) the notions of duality and dual problems, arising from the $S(F,G)$ description are examined. This topic provides the motivation for the material in chapter (6).

The fourth chapter generalises the Segrè-Weyr characteristic theory and the generalised null spaces results from the standard eigenvalue-eigenvector problem defined by a regular pencil. For every generalised eigenvalue ($|\lambda F-G|=0$), a special sequence of the "Piecewise Arithmentic Progression" (PAP) type is defined; the analysis of PAP sequences, leads to the derivation of the Segrè characteristic of (F,G) at $s=\lambda$, which in turn, provides a number theoretic definition of e.d., independetly from the algebraic one. A method for constructing the Weierstrass canonical form is suggested, which makes use of Ferrer's diagrams for the analysis of PAP sequences and avoids the use of special transformations. The geometry of a regular pencil is investigated by defining the notion of the generalised null space at $s=\lambda$ and then by investigating its structural properties. The structure of basis matrices for the null space of Toeplitz matrices reveals the complete geometric dimension of e.d.; a systematic procedure for computing independent Jordan type chains is given, which in turn provides an alternative technique for computing the Weierstrass form. The results of the chapter reveal the complete geometric and number theoretic dimension of e.d. and provide alternative means for their definition/computation.

Chapter (5) is an extension of the Segrè-Weyr theory to the case of singular pencils. It is shown that the number theoretic and geometric results developed for regular pencils may extended for the case of column, row minimal indices (c.m.i.), (r.m.i.) of a singular pencil. In fact, the set of c.m.i. (r.m.i.) may be defined by analysing the properties of appropriate piecewise Arithmetic Progression sequences; these results demonstrate the unity between the S.E. invariants and provide the number theoretic dimension of c.m.i. (r.m.i.) is associated with the structure of the maximal annihilating spaces defined on a pair (F,G). The module structures of the pair (F,G) are examined and a systematic procedure for constructing minimal polynomial bases, based on standard linear algebra, is suggested. The results provide a thorough description of the number theoretic, geometric and algebraic properties of c.m.i., r.m.i.

The sixth chapter develops a general theory of algebraic invariants of matrix pencils, or ordered pairs, under Bilinear Strict Equivalence (B.S.E). Starting from the invariance property of c.m.i., r.m.i. and the covariance property of e.d. a complete set of invariants under B.S.E. is defined. These invariants are defined by the real, complex lists (degrees of all e.d., or all Segrè characteristics) and vectors defined as Plücker vectors and canonical Grassmann vectors. A fundamental part of the analysis is the determination of a complete set of invariants of homogeneous binary polynomials under projective equivalence; this problem is shown to be equivalent to the general linear mapping problem on a Riemann sphere. Finally, the effect of B.E. on stability of e.d. is examined and the role of B.E. transformations on assigning different properties to matrix pencils is discussed. These results in this chapter provide the tools for the relativistic classification of system properties, construction of dual problems and investigating the existence of pencils with "good" numerical analysis characteristics.

Chapter (7) deals with the number theoretic, geometric and dynamic apsects of a general pencil and it is derived into four main parts. Section (7.2) extends the theory of PAP sequences characterising the sets of elementary divisors (e.d.) to the general case of singular pencils. These results together with the PAP sequences characterisation of c.m.i, and r.m.i, provide the means for a number theoretic definition of the set of S.E, invariants on the pair (F,G), as well as a procedure for finding the Kronecker form without use of special transformations; The basic element of the procedure is the singular value decomposition which is used for the computation of the different types of PAP sequences. Section (7.3) deals with the geometric properties of the subspaces of the domain of (F,G). The key tool is the restriction pencil $(F,G)/\mathcal{V}$ with $I_{\mathcal{V}}$ set of S.E. invariants. The geometry of simple invariant subspaces (characterised by one type only of S.E. invariants) is examined first; the stucture of such subspaces leads to the definition of (F,G)-, (G,F)-, and complete (F,G)-invariant subspaces. A detailed account of the structure of such subspaces is given and these notions of invariant subspaces are extended to those of partitioned-(F,G)-invariant subspaces (p.-(F,G)-i.s.), extended (F,G)-, (G,F)-, complete-(F,G)-invariant subspaces. The geometry and spectrum properties of such subspaces are examined in detail. Finally, the relationships of abstract subspace algorithms defined on (F,G) pair with the different notions of invariant subspaces is discussed. This latter study leads to a characterisation and computation of the various types of supremal invariant subspaces of the domain of (F,G). The results of this section establish a complete geometric theory for matrix pencils, which may be presented on ordered pairs (F,G) and independently from the underlying algebra.

In section (7.4) the theory of invariant realisations [Kar. & Hay. -1,2] is further developed. The theory is presented on ordered triples $(F,G;\mathcal{V})$, where is a subspace of the domain of (F,G); these results allow the interpretation of the geometric and dynamic results derived on $\mathcal{S}(F,G)$ descriptions, or matrix pencils, back to the theory of linear systems. The essence of the theory is the definition of another triple $(F',G';\mathcal{V}')$, where (F',G') is entirely right singular (characterised only by c.m.i.), such that the restriction pencils $(F,G)/\mathcal{V}$, $(F',G')/\mathcal{V}'$ are S.E. The final section (7.5), deals with the properties of the solution space of the $\mathcal{S}(F,G)$ descriptions and the dynamic characterisation of the invariant subspaces of (F,G). The solution space is characterised by defining the initial and redundancy spaces of $\mathcal{S}(F,G)$; the initial space is further characterised by defining the regular (uniqueness of solutions) and non regular (no uniqueness of solutions) subspaces. A further classification of the regular initial space is given in terms of the C^{∞}, distributional nature of the solutions. These notions of initial subspaces are related to the invariant subspaces of the pair (F,G). The dynamic characterisation of invariant subspaces is done in terms of two fundamental properties; the holdability and the reachability properties. According to the nature of solutions we distinguish the familes of C^{∞}-, distributionally-holding subspaces and C^{∞}- , distributionally reachability subspaces; these notions, defined on $\mathcal{S}(F,G)$ systems, are extensions of the notions of (A,B)-invariant, almost (A,B)-invariant, controllability, almost controllability subspaces of the geometric theory. The different types of invariant subspaces of (F,G) are classified according to the above four properties. Finally, the parametrisation of the solution is discussed; this is achieved by using the notion of invariant forced realisation.

Chapter (8) provides a framework for discussing the topological (properties under uncertainty) aspects of pairs (F,G) and a "space-frequency" relativistic classification of the geometric and dynamic properties of $\mathcal{S}(F,G)$ systems. Different types of convinient (as far as description of uncertainty) metric topologies are introduced. The new metrics are related to already known ones; thus, the links with standard perturbation results are established. The notion of deflating subspace [Stew. -1], which plays an important role in perturbation theory, is shown to be equivalent to the notion of elementary divisor type subspace discussed in section (7.3). Finally, the results of chapter (6) are used to classify the geometric and dynamic properties of $S(F,G)$ systems to those which depend on frequency coordinate transformations and those which are invariant under such transformations. The conclusions and future research extensions of this work are given in chapter (9).

CHAPTER 2:

Mathematical Background

CHAPTER 2: MATHEMATICAL BACKGROUND

In this chapter we give a brief account of some important mathematical topics which form an essential background to the topics discussed in the folowing Chapters.

In Section (2.1) we review some of the fundamental concepts and results on invariants and canonical forms defined on equivalence classes. In Section (2.2) a brief account of results from the theory of polynomial matrices and rational vector spaces is given. Particular emphasis is given to the theory of polynomial minimal bases, which in the context of matrix pencils plays a crucial role, especially in Chapter (5). The background on Exterior algebra is needed for the developement of the theory of invariants of matrix pencils under Bilinear strict equivalence, presented in Chapter (6), as well as for the developement of a new angle metric for ordered pairs (F,G), or matrix pencils given in the last Chapter. In Section (2.4) a brief review of the results on Segrè characteristics of the standard eigenvalue problem is given; this review is needed as a background to Chapters (4) and (5), where the Segre' theory is extended to the generalised eigenvalue-eigenvector problem. A brief account of the classical matrix pencil theory is given in Section (2.5). Finally, a summary of essential topological notions is given in Section (2.6).

2.1 Equivalence Relations, Invariants and Canonical forms [Bir. & MacL. -1]

We shall begin by introducing some definitions and results which will be used later in this thesis.

Definition (2.1): A relation from a set X to a set Y is a subset of $X\mathrm{x}Y$ where $X\mathrm{x}Y$ is the set of all ordered pairs of the form (x,y) where $x\in X$, $y\in Y$.

Definition (2.2): A relation R on X ($R\subset X\mathrm{x}X$) is called an equivalence relation if it satisfies the following three conditions

(i) $\forall\, x\in X$, $(x,x)\in R$ (reflexive)

(ii) If $(x,x')\in R$ then $(x',x)\in R$ (symmetric)

(iii) If $(x,x')\in R$ and $(x',x'')\in R$ then $(x,x'')\in R$ (transitive)

Definition (2.3): Let X be a set ($X\neq\emptyset$) and E be an equivalence relation on X. Let $x\in X$; then the equivalence class or orbit of x is befined by:

$$E(x)=\{y:\ y\in X \text{ and } (x,y)\in E\}\ , \tag{2.1}$$

The set of all equivalence classes is called the quotient set or orbit set and it is denoted by X/E

Theorem (2.1): If R is an equivalence relation on X, then the family of all equivalence classes form a partition of X, i.e.

$$X=R(x_1)\cup R(x_2)\cup\ldots\cup R(x_i)\cup\ldots\quad ,\qquad \underset{i\neq j}{R(x_i)\cap R(x_j)=\emptyset}$$

Definition (2.4): A system of distrinct representatives for R, is a set T, $T\subseteq X$ that contains precisely one element from each of the equivalence classes.

Definition (2.5): Let X,T are sets, E an equivalence relation defined on X. A function $f:X\to T$ is called an invariant of E, when xEy implies $f(x)=f(y)$. $f:X\to T$ is called a complete invariant for E, when $f(x)=f(y)$ implies xEy.

*(A complete invariant defines a one to one correspondence between the equivalence classes $E(x)$ and the image of f).

Definition (2.6): A set of invariants $\{f_i, f_i:X\to T_i\ \ i=1,2,\ldots,k\}$ is a complete set for E, if the map f defined by:

$$f:X\to \underset{i=1}{\overset{k}{\times}} T_i:\ x\to f(x)\triangleq(f_1(x),f_2(x),\ldots,f_k(x)) \tag{2.2}$$

is a complete invariant for E on X.

Let $f:X\to \underset{i=1}{\overset{k}{\times}} T_i$ is a complete invariant for E on X. Then $f_i(x)$, $i=1,2,\ldots,k$ characterise uniquely $E(x)$. If we specialize the invariant f such that its image $T\subset X$ we define a canonical element or a canonical form.

Definiition (2.7): A set of canonical forms for $\mathcal{E}$ equivalence on $\mathcal{X}$ is a subset $\mathcal{C}$ of $\mathcal{X}$ such that $\forall\ x\in\mathcal{X}$ there exists a unique $c\in\mathcal{C}$ for which $x\mathcal{E}c$.

Indeed, if f is a complete invariant and $f(\mathcal{X})=\mathcal{C}$, then for any $x\in\mathcal{X}$ and c_1, $c_2\in\mathcal{C}$, $x\mathcal{E}c_1$ and $x\mathcal{E}c_2$ implies $f(x)=f(c_1)=f(c_2)=c_1=c_2=c$ by the invariance property. By completeness, we have that for any $c\in\mathcal{C}$, if $f(x_1)=c$ and $f(x_2)=c$, then $x_1\mathcal{E}x_2$.

Thus, $c=f(x)$ is a unique member of $\mathcal{E}(x)$ $\forall\ x\in\mathcal{X}$.

The values $f_i(x)$ are often called a complete set of invariants. □

2.2. Polynomial matrices, Rational Vector Spaces and Smith form

2.2.1. $\mathbb{R}[s]$-unimodular equivalence, Smith form. [Gant. 1]

Let $\mathbb{R}^{p\times q}[s]$ the set of pxq matrices from $\mathbb{R}[s]$; an equivalence relation on this set may be defined as follows:

Definition (2.8): Let $M_1(s),M_2(s)$ $\mathbb{R}^{p\times q}[s]$. $M_1(s),M_2(s)$ are said to be $\mathbb{R}[s]$-equivalent if

$$M_2(s)=R(s)M_1(s)Q(s), \tag{2.3}$$

where $R(s)\in\mathbb{R}^{p\times p}[s]$, $Q(s)\in\mathbb{R}^{q\times q}[s]$, $|R(s)|$, $|Q(s)|\in\mathbb{R}-\{0\}$.

The matrices $R(s)$, $Q(s)$ with detemrinants in $\mathbb{R}-\{0\}$ are known as $\mathbb{R}[s]$-unimodular matrices. If $R(s)=I_p$ and $Q(s)$ $\mathbb{R}(s)$-unimodular, then $M_1(s)$, $M_2(s)$ are called $\mathbb{R}[s]$-right-equivalent and if $Q(s)=I_q$, and $R(s)$ $\mathbb{R}[s]$-unimodular, then they are called $\mathbb{R}[s]$-left-equivalent. It is well known that the relations defined above are equivalence relations and a complete set of invariants may be defined, as well as canonical forms. The $\mathbb{R}[s]$-equivalence class of $M(s)\in\mathbb{R}^{p\times q}[s]$ will be denoted by $\mathcal{E}_{\mathbb{R}[s]}(M)$ and the right, left $\mathbb{R}[s]$-equivalence classes by $\mathcal{E}^r_{\mathbb{R}[s]}(M)$, $\mathcal{E}^l_{\mathbb{R}[S]}(M)$ respectively.

Theorem (2.1): Let $M(s)\in\mathbb{R}^{pxq}[s]$.

(i) There exist unimodular matrices $L(s)\in\mathbb{R}^{pxp}[s]$ (obtained by e.r.o.'s.) and $R(s)\in\mathbb{R}^{qxq}[s]$ (obtained by e.c.o.'s.) such that

$$L(s)M(s)R(s)=S(s)=\begin{array}{cl} \left[\begin{array}{ccccc:c} \lambda_1(s) & & & & & \\ & \cdot & & 0 & & \\ & & \lambda_i(s) & & & 0 \\ 0 & & & \cdot & & \\ & & & & \lambda_r(s) & \\ \hdashline & & 0 & & & 0 \end{array}\right] & \begin{matrix} r \\ \\ \\ \\ p-r \end{matrix} \\ \qquad\qquad r \qquad\qquad\qquad q-r & \end{array} \in \mathbb{R}^{pxq}[s] \qquad (2.4)$$

where $S(s)$ is called the Smith form of $M(s)$ and

(a) $r=\max\limits_{s\in\mathbb{R}}\{\text{rank } M(s)\} \leqq \min(p,q)$

(b) the polynomials $\lambda_i(s)\in\mathbb{R}[s]$, $i\in\underset{\sim}{r}$ called invariant polynomials of $M(s)$, are monic, uniquely defined by $M(s)$ and satisfy the division property

$\lambda_i(s)/\lambda_{i+1}(s)\ \forall\ i\in\underset{\sim}{r-1}$.

(ii) The set $\{\lambda_i(s),\ i\in\underset{\sim}{r}\}$ of invariant polynomials defined above forms a complete set of invariants for $E_{\mathbb{R}[s]}$-equivalence and $S(s)$ is a canonical form for $E_{\mathbb{R}[s]}(M)$.

Moreover, consider the polynomials $\Delta_i(s)\in\mathbb{R}[s]$, $i=0,1,\ldots,r$ given by

$\Delta_0(s)\equiv 1$, $\Delta_i(s)$:=the monic g.c.d. of all ixi minors of $M(s)$.

These polynomials are called the determinantal divisors of $M(s)$. They are related to the invariant polynomials of $M(s)$ by the Smith algorithm.

$$\text{i.e.}\quad \lambda_i(s)=\frac{\Delta_i(s)}{\Delta_{i-1}(s)},\quad i\in\underset{\sim}{r} \qquad (2.5)$$

If $\lambda_i(s)=(s-s_{i,1})^{\alpha_{i1}}\ldots(s-s_{i,k_i})^{\alpha_{ik_i}}$, $i\in\underset{\sim}{r}$ is the unique factorisation over ¢ of $\lambda_i(s)$, where $s_{i,1},\ldots,s_{i,k_i}\in\mathbb{C}$ and $\alpha_{il},{}_{i2},\ldots,\alpha_{ik_i}\in Z^+$, then, the elements of the set $\{(s-s_{ij})^{\alpha_{ij}},\ j=1,2,\ldots k_i,\ i=1,2,\ldots,k\}$ are called the elementary divisors of $M(s)$. An elementary divisor is said to be linear or non linear according as $\alpha_{ij}=1$ or $\alpha_{ij}>1$.

Remark (2.1): Some different elementary divisors may contain the same polynomial $(s-s_o)^{\alpha}$, (this happens, for example, in case $\lambda_i(s)=\lambda_{i+1}(s)$ for some i). The total number of elementary divisors of M(s) is therefore $\sum_{i=1}^{r} k_i$. □

Remark (2.2): Let $M(s)\in\mathbb{R}^{nxn}[s]$ and $\det M(s)\not\equiv 0$. Then the sum $\sum_{i=1}^{r}\sum_{j=1}^{k_i}\alpha_{ij}$ of degrees of its elementary divisors $(s-s_{ij})^{\alpha_{ij}}$ coincides with the degree of detM(s).

Note that the knowledge of the elementary divisors of M(s) and of the number r of invariant polynomials $\lambda_1(s),\ldots,\lambda_r(s)$ is sufficient to construct $\lambda_1(s)$, $\ldots,\lambda_r(s)$. In this construction we use the fact that $\lambda_i(s)$ is divisible $\lambda_{i-1}(s)$. Let $s_1,s_2,\ldots,s_p$ be all the different numbers from $\mathbb{C}$ which appear in the elementary divisors, and let that $(s-s_i)^{\alpha_{i,1}},\ldots,(s-s_i)^{\alpha_{i,k_i}}$, $(i\in\underline{p})$ be the elementary divisors containing the number s_i, and ordered in the descending order of the degrees $\alpha_{i1}\geqq\cdots\geqq\alpha_{ik_i}>0$. Clearly, the number r of invariant polynomials must be greater than equal to max $\{k_i,\ldots,k_p\}$.

Under this condition, the invariant polynomials $\lambda_1(s),\ldots,\lambda_r(s)$ are given by the formulas:

$$\lambda_j(s)=\prod_{i=1}^{p}(s-s_i)^{\alpha_{i,r+1-j}},\quad j\in\underset{\sim}{r} \tag{2.6}$$

where we put $(s-s_i)^{\alpha_{i,j}}=1$ for $j>k_i$ □

2.2.2. Rational matrices, M.F.Ds, coprimeness [Kail., 1]

Consider the rational matrix $G(s)\in\mathbb{R}^{m\times\ell}(s)$, $\text{rank}_{\mathbb{R}(s)}\{G(s)\}=\min\{m,\ell\}$. It is then well known that G(s) can always be factored (in a non-unique way) as

$$G(s)=D_L^{-1}(s)N_L(s)=N_R(s)D_R^{-1}(s) \tag{2.7}$$

where $N_L(s)$, $N_R(s)\in\mathbb{R}^{m\times\ell}[s]$, $D_L(s)\in\mathbb{R}^{m\times m}[s]$, $D_R(s)\in\mathbb{R}^{\ell\times\ell}[s]$ with det $D_L(s)$, det $D_R(s)\not\equiv 0$. The pair $\{D_R(s), N_R(s)\}$ ($\{D_L(s), N_L(s)\}$) is called a right (left) matrix fraction description (MFD) of the rational matrix G(s).

The above definition, and the other results to be described later, show that matrix fraction descriptions provide a natural generalisation of the scalar rational function representation of a scalar rational function.

Furthermore, descriptions (2.7) establish the links between rational matrix theory and polynomial matrix theory.

Definition (2.9): A square polynomial matrix $Q(s) \in \mathbb{R}^{q\times q}[s]$ is said to be a right divisor (R.D.) of a polynomial matrix $M(s) \in \mathbb{R}^{p\times q}[s]$, with $p \geq q$, if and only if there exists a polynomial matrix $M_1(s) \in \mathbb{R}^{p\times q}[s]$, such that

$$M(s) = M_1(s)Q(s) \qquad (2.8)$$

Let $G(s)$ be a R.D. of $M(s)$. Then $Q_G(s)$ is said to be a greatest right divisor (G.R.D.) of $M(s)$ if and only if $\deg\{\det Q_G(s)\} \geqq \deg\{\det Q(s)\}$ for every R.D. $Q(s)$ of $M(s)$.

□

Remark (2.3): Greatest right divisors of polynomial matrices are not unique. They differ only by unimodular (left) factors.

Definition (2.10): A polynomial matrix $M(s) \in \mathbb{R}^{p\times q}[s]$, $p \geqq q$, $\text{rank}_{\mathbb{R}(s)}\{M(s)\}=q$ is said to be irreducible or least degree if and only if one of the following equivalent conditions is satisfied:

(i) all the G.R.D. of $M(s)$ are unimodular matrices;

(ii) the Smith form of $M(s)$ is $\begin{bmatrix} I_q \\ 0 \end{bmatrix}$;

(iii) the greatest common divisor of all q-order minors of $M(s)$ is 1;

(iv) $\text{rank}_{\mathbb{R}(s)}\{M(s)\}=q$, for every $s \in \mathbb{C}$.

□

Definition (2.11): A square polynomial matrix $Q(s) \in \mathbb{R}^{q\times q}[s]$ is said to be a greatest common right divisor (G.C.R.D.) of the two polynomial matrices $M_1(s) \in \mathbb{R}^{p\times q}[s]$, $M_2(s) \in \mathbb{R}^{m\times q}[s]$ if and only if $Q(s)$ satisfies the following properties:

(i) $Q(s)$ is a common right divisor of $M_1(s)$ and $M_2(s)$;

(ii) if $Q'(s) \in \mathbb{R}^{qxq}[s]$ is any other common right divisor of $M_1(s)$ and $M_2(s)$, then $Q'(s)$ is a right divisor of $Q(s)$, or in other words $\deg\{\det\{Q(s)\}\} \geq \deg\{\det\{Q'(s)\}\}$.

□

Remark (2.4): Greatest common right divisors of two polynomial matrices are not unique. They differ only by unimodular (left) factors.

□

Definition (2.12): Two polynomial matrices $M_1(s) \in \mathbb{R}^{pxq}[s]$, $M_2(s) \in \mathbb{R}^{mxq}[s]$ with $\text{rank}_{\mathbb{R}(s)} \{\begin{bmatrix} M_1(s) \\ M_2(s) \end{bmatrix}\}=q$ are said to be relatively right prime or right coprime if and only of one of the following equivalent conditions is satisfied:

(i) all the G.C.R.D. of $M_1(s)$ and $M_2(s)$ are unimodular matrices;

(ii) the Smith form of $\begin{bmatrix} M_1(s) \\ M_2(s) \end{bmatrix}$ is $\begin{bmatrix} I_q \\ 0 \end{bmatrix}$;

(iii) the greatest common divisor of all q-order minors of $\begin{bmatrix} M_1(s) \\ M_2(s) \end{bmatrix}$ is 1;

(iv) $\text{rank}_{\mathbb{R}(s)} \{\begin{bmatrix} M_1(s) \\ M_2(s) \end{bmatrix}\}=q$, for every $s \in C$.

□

Remark (2.5): Left divisors (L.D.), Greatest left divisors (G.L.D) and Greatest common left divisors (G.C.L.D.) can be defined with the obvious changes. For convinience, we shall henceforth talk only of right divisors.

□

Remark (2.6): A right MFD (left MFD) $\{D_R(s), N_R(s)\}$ ($\{D_L(s), N_L(s)\}$) of a transfer function matrix $G(s)$ is called a right coprime MFD (a left coprime MFD) if and only if the matrixes $D_R(s)$, $N_R(s)$ ($D_L(s)$, $N_L(s)$) are tight coprime (left coprime).

Let now $M(s) \in \mathbb{R}^{pxq}[s]$, $p \geq q$ be a polynomial matrix with $\text{rank}_{\mathbb{R}(s)}\{M(s)\}=q$ and let us write it in terms of its q column polynomial vectors as

$$M(s) = [\underline{m}_1(s),\dots,\underline{m}_q(s)] \tag{2.9a}$$

where

$$\underline{m}_i(s) = [m_{1i}(s),\dots,m_{pi}(s)] \quad i=1,\dots,q \tag{2.9b}$$

Definition (2.13): The degree of the polynomial vector $\underline{m}_i(s)$ is the highest degree occuring among the degrees of its polynomial elements $m_{ji}(s)$, i.e.

$$\deg \underline{m}_i(s) = \max_{j=1,\dots,p} \{\deg m_{ji}(s)\} \qquad i=1,\dots,q \tag{2.10}$$

$\square$

Definition (2.14) [Rosenbrock -1]: The complexity of M(s) is the sum of the degrees of its column polynomial vectors, i.e.

$$c = \sum_{i=1}^{q} \deg\{\underline{m}_i(s)\} \tag{2.11}$$

$\square$

Definition (2.15) [Rosenbrock -1]: The degree d of M(s) is the highest degree occuring among the degrees of all its q-order minors.

Since a q-order minor of M(s) is a sum of products of polynomials one from each column, the maximum degree occuring among all the q-order minors of M(s), i.e. its degree d cannot exceed its complexity c, i.e. we have [Rosenbrock -1] $c \geqq d$. Let now that $g_i = \deg \underline{m}_i(s)$, $i=1,\dots,q$, and write

$$\underline{m}_i(s) = \underline{m}_i^o + \underline{m}_i^1 s + \dots + \underline{m}_i^{g_i} s^{g_i} = \sum_{k=0}^{g_i} \underline{m}_i^k s^k, \quad i=1,\dots,q \tag{2.12}$$

Then M(s) can be written as

$$M(s) = [\underline{m}_1(s),\dots,\underline{m}_q(s)] = [\underline{m}_1^{g_1},\dots,\underline{m}_q^{g_q}] \begin{bmatrix} s^{g_1} & & 0 \\ & \ddots & \\ 0 & & s^{g_q} \end{bmatrix} + M_b Z(s) \tag{2.13}$$

where $M_b \in \mathbb{R}^{p\times c}$ $(c = \sum_{i=1}^{q} g_i)$, and

$$Z(s) = \begin{bmatrix} 1 & & 0 \\ s & & \\ \vdots & & \\ s^{g_1-1} & & \\ & \ddots & \\ & & 1 \\ & & s \\ & & \vdots \\ 0 & & s^{g_q-1} \end{bmatrix} \in \mathbb{R}^{c\times q}[s] \tag{2.14}$$

The matrix $[\underline{m}_1^{g_1},\dots,\underline{m}_q^{g_q}]=M_a \in \mathbb{R}^{p\times q}$ is called the highest (column) degree coefficient matrix of M(s).

□

Definition (2.16): A polynomial matrix $M(s) \in \mathbb{R}^{p\times q}[s]$ is said to be column proper or column reduced if the matrix M_a has full rank q.

□

Proposition (2.1): A polynomial matrix $M(s) \in \mathbb{R}^{p\times q}[s]$ is a column reduced if its complexity c is equal to its degree d.

□

Proposition (2.2): Let $M(s) \in \mathbb{R}^{p\times q}[s]$ be a polynomial matrix which is not column reduced. Then there always exists a unimodular matrix $U(s) \in \mathbb{R}^{q\times q}[s]$, $\det\{U(s)\} \in \mathbb{R}-\{0\}$, such that the polynomial matrix $M'(s)=M(s)U(s)$ is column reduced.

□

The Algebraic Structure of Rational Vector Spaces

Let $G(s) \in \mathbb{R}^{m\times \ell}[s]$, $m\geq\ell$, $\text{rank}_{\mathbb{R}(s)}\{G(s)\}=\ell$ be a transfer function matrix. Let us also denote by $\mathcal{V}_G$ the set of all linear combinations of the columns of G(s) with multipliers in $\mathbb{R}(s)$, i.e.

$$\text{if } G(s) = [\underline{g}_1(s),\dots,\underline{g}_\ell(s)], \text{ then } \mathcal{V}_G = \text{span}_{\mathbb{R}(s)}\{\underline{g}_1(s),\dots,\underline{g}_\ell(s)\} \quad (2.15)$$

Clearly $\mathcal{V}_G$ is a linear vector space over $\mathbb{R}(s)$ and $\dim\mathcal{V}_G=\ell$, and it is called the rational vector space generated by G(s).

From any rational basis G(s) of $\mathcal{V}_G$ we can generate a polynomial basis of $\mathcal{V}_G$ by means of a right MFD of G(s), i.e. if $G(s)=N(s)D^{-1}(s)$ with $N(s) \in \mathbb{R}^{m\times\ell}[s]$, $D(s) \in \mathbb{R}^{\ell\times\ell}[s]$, $\det\{D(s)\}\neq 0$, then clearly the columns of N(s) define a polynomial basis of $\mathcal{V}_G$. More precisely, if $N(s)=[\underline{n}_1(s),\dots,\underline{n}_\ell(s)]$ then

$$\text{span}_{\mathbb{R}(s)}\{\underline{n}_1(s),\dots,\underline{n}_\ell(s)\} = \mathcal{V}_G$$

and (2.16)

$$\text{span}_{\mathbb{R}(s)}\{\underline{n}_1(s),\dots,\underline{n}_\ell(s)\} = \mathcal{M}_N$$

where $\mathcal{M}_N$ denotes the set of all linear combinations of the columns of N(s) with multipliers in $\mathbb{R}[s]$. The set $\mathcal{M}_N$ is a free $\mathbb{R}[s]$-module [Bir. -1] and it is called the <u>polynomial module generated by N(s)</u>.

□

<u>Proposition (2.3)</u>: Let $\mathcal{M}_{N_1}\mathcal{M}_{N_2}$ be the polynomial modules generated by the polynomial matrices $N_1(s)$, $N_2(s) \in \mathbb{R}^{m\times\ell}[s]$, with $\text{rank}_{\mathbb{R}(s)}\{N_1(s)\}=\text{rank}_{\mathbb{R}(s)}\{N_2(s)\}=\ell$. If $N_1(s)=N_2(s)Q(s)$, where $Q(s) \in \mathbb{R}^{\ell\times\ell}[s]$, $\det\{Q(s)\}\not\equiv 0$, then

$$\mathcal{M}_{N_1} \subseteq \mathcal{M}_{N_2} \tag{2.17}$$

□

<u>Proposition (2.4)</u>: Let $N_1(s)$, $N_2(s) \in \mathbb{R}^{m\times\ell}[s]$ be two polynomial bases of the same polynomial module $\mathcal{M}_N$. Then, there exists a unimodular matrix $Q(s) \in \mathbb{R}^{\ell\times\ell}[s]$, $\det\{Q(s)\}=c \in \mathbb{R}-\{0\}$ such that

$$N_1(s) = N_2(s)Q(s) \tag{2.18}$$

Thus unimodular matrices represent coordinate transformations of a polynomial module.

<u>Proposition (2.5)</u>: Let $N(s) \in \mathbb{R}^{m\times\ell}[s]$ be a basis of the polynomial module $\mathcal{M}_N$. Then the degree of N(s) is an invariant of $\mathcal{M}_N$, or in other words if $N_1(s) \in \mathbb{R}^{m\times\ell}[s]$ is any other basis of $\mathcal{M}_N$ then $\deg\{N(s)\}=\deg\{N_1(s)\}$.

□

<u>Proposition (2.6)</u>: Let $N_1(s)$, $N_2(s) \in \mathbb{R}^{m\times\ell}[s]$, $m\geq\ell$, $\text{rank}_{\mathbb{R}(s)}\{N_1(s)\}=\ell$, $\text{rank}_{\mathbb{R}(s)}\{N_2(s)\}=\ell$ and let $d_1=\deg\{N_1(s)\}$, $d_2=\deg\{N_2(s)\}$. If

$$N_1(s) = N_2(s)Q(s),\ Q(s)\in\mathbb{R}^{\ell\times\ell}[s],\quad \deg\{\det Q(s)\} = q\geq 1 \tag{2.19}$$

then

$$\text{(i)}\quad d_1 = d_2 + q \tag{2.20}$$

$$\text{(ii)}\quad \mathcal{M}_{N_1} \subset \mathcal{M}_{N_2} \tag{2.21}$$

where $\mathcal{M}_{N_1}$, $\mathcal{M}_{N_2}$ are the polynomial modules generated by the polynomial matrices $N_1(s)$, $N_2(s)$, respectively.

Clearly, Eqn. (2.19) represents the extraction of a right divisor $Q(s)$ of the polynomial matrix $N_1(s)$. This observation leads us to the following conclusions:

Let $N_1(s) \in \mathbb{R}^{m\times\ell}[s]$, $m\geq\ell$, $\text{rank}_{\mathbb{R}(s)}\{N(s)\}=\ell$ be a polynomial matrix which can be written in terms of its columns as $N_1(s)=[\underline{n}_1^1(s),\ldots,\underline{n}_\ell^1(s)]$. Let us assume that $N_1(s)$ is not irreducible and let

$$V = \text{span}_{\mathbb{R}(s)}\{\underline{n}_1^1(s),\ldots,\underline{n}_\ell^1(s)\},\ M_{N_1} = \text{span}_{\mathbb{R}[s]}\{\underline{n}_1^1(s),\ldots,\underline{n}_\ell^1(s)\} \quad (2.22)$$

be the rational vector space V and the polynomial module M_{N_1} spanned by its columns. Then if $Q_i(s)$, $i=1,2,\ldots$ are right divisors of $N_1(s)$, i.e.

$$N_1(s) = N_{i+1}(s)Q_i(s), \quad i=1,2,\ldots \quad (2.23)$$

and the $\deg\{\det Q_i(s)\}=q_i\geq 1$ are such that $q_1\leq q_2\leq q_3\ldots$ we will have that

$$M_{N_1} \subset M_{N_2} \subset M_{N_3} \subset \ldots \quad (2.24)$$

and

$$\deg(N_1(s)) \geq \deg(N_2(s)) \geq \deg(N_3(s)) \geq \ldots \quad (2.25)$$

Moreover, if $Q_G(s)$ is a greatest right divisor of $N_1(s)$ so that $N_1(s)= N(s)Q_G(s)$, then

$$M_{N_1} \subset M_N \qquad \text{and} \qquad \deg(N_1(s)) \geq \deg(N(s)) \quad (2.26)$$

The polynomial module M_N is the <u>maximal submodule</u> of the rational vector V and all its polynomial bases are least degree or irreducible polynomial matrices. In other words, if we consider the set of all polynomial vectors in V then this set coincides with the module M_N defined above.

Clearly, although the ascending chain of modules which is defines by Eqn. (2.24) is not unique (it depends on the choice of $Q_i(s)$) the maximal module M_N is uniquely defined.

□

Definition (2.17) [For. -1]: A polynomial matrix $N(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$, $m \geq \ell$ and $\mathrm{rank}_{\mathbb{R}(s)}\{N(s)\}=\ell$ is said to be a minimal basis of the rational vector space $\mathcal{V}$ spanned by the columns of a polynomial matrix $N_1(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$, $m\geq\ell$, $\mathrm{rank}_{\mathbb{R}(s)}\{N_1(s)\}=\ell$, if and only if $N_1(s)=N(s)Q_G(s)$ where $G_G(s)$ is a greatest right divisor of $N_1(s)$ and $N(s)$ has the following properties:

(i) N(s) is least degree;

(ii) N(s) is column reduced. □

Remark (2.7): Let $N_1(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$, $m\geq\ell$, $\mathrm{rank}_{\mathbb{R}(s)}\{N_1(s)\}=\ell$. If $N(s)$, $N^*(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$ are two minimal bases of the rational vector space $\mathcal{V}$ spanned by the columns of N(s), then

$$N(s) = N^*(s)Q(s) \tag{2.27}$$

where $Q(s) \in \mathbb{R}^{\ell\mathrm{x}\ell}[s]$ is a unimodular matrix. □

Remark (2.8): Let $\underline{x}(s)$ be a polynomial vector of the rational vector space $\mathcal{V}$ spanned by the columns of $N_1(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$, $m\geq\ell$, $\mathrm{rank}_{\mathbb{R}(s)}\{N_1(s)\}=\ell$ and let N(s) be a minimal basis of $\mathcal{V}$. Then $\underline{x}(s)$ can be expressed as a polynomial combination of the columns of N(s).

Given in general, a $G(s) \in \mathbb{R}^{m\mathrm{x}\ell}[s]$, $m\geq\ell$, $\mathrm{rank}_{\mathbb{R}(s)}\{G(s)\}=\ell$, then Forney - describes a way of computing a minimal basis for the rational vector space spanned by its columns (i.e. $\mathcal{V}_G$). He then shows that the column degrees $\delta_i = \mathrm{deg}\underline{n}_i(s)$, $i=1,\ldots,\ell$ of a minimal basis $N(s)=[\underline{n}_1(s),\ldots,\underline{n}_\ell(s)]$ are the same (i.e. invariant) for every minimal basis of $\mathcal{V}_G$, i.e. δ_i, $i=1,\ldots,\ell$ characterise $\mathcal{V}_G$. Forney calls these degrees the "invariant dynamic indices" of $\mathcal{V}_G$ and their sum

$$\delta = \sum_{i=1}^{\ell} \delta_i \tag{2.28}$$

the "invariant dynamic order" of $\mathcal{V}_G$. Clearly, the invariant dynamical order is the complexity of N(s) and since N(s) is a minimal basis it is equal to the degree of N(s). We have to note that the invariant dynamical indices and the invariant dynamical order do not characterise V_G uniquely, i.e. they are not complete invariants.

□

2.3. Background from Exterior Algebra [Gre. -1]

Let V be an arbitrary vector space and $p \geq 2$ be an integer. Then a vector space $\Lambda^p V$ together with a skew symmetric p-linear map

$$\Lambda^p : \overset{p}{\underset{1}{x}} V \to \Lambda^p V \tag{2.29}$$

is called a p-th exterior power of V if the following conditions are satisfied:

(i) The vectors $\Lambda^p(\underline{x}_1, \ldots, \underline{x}_p)$ generate $\Lambda^p V$.

(ii) If ψ is any skew symmetric p-linear map of $\overset{p}{\underset{1}{x}} V$ into an arbitrary vector space U, then there exists a linear map $f : \Lambda^p V \to U$ such that $\psi = f \circ \Lambda^p$.

It is proved that conditions (i) and (ii) are equivalent to the following condition

(iii) If ψ is any skew symmetric p-linear map of $\overset{p}{\underset{1}{x}} V$ into a vector space U, then there exists a unique linear map $f : \Lambda^p V \to U$ such that $\psi = f \circ \Lambda^p$.

The above conditions may be summarized by the following commutative diagram:

$$\begin{tikzcd} \overset{p}{\underset{1}{x}} V \arrow[r, "\Lambda^p"] \arrow[dr, "\psi"'] & \Lambda^p V \arrow[d, "f"] \\ & U \end{tikzcd}$$

where:

Λ^p : p-linear skew symmetric

ψ : p-linear skew symmetric

f : linear

Figure (2.1)

The elements of $\Lambda^p V$ are called p-vectors. A p-vector of the form $\Lambda^p(\underline{x}_1, \ldots, \underline{x}_p)$ is called decomposable, and it is denoted by $\underline{x}_1 \wedge \ldots \wedge \underline{x}_p$. Condition (i) states that $\Lambda^p V$ is generated by its decomposable elements.

The skew symmetric property of the p-linear map Λ^p implies that for every permutation $\sigma \in S_p$

$$\underline{x}_{\sigma(1)} \wedge \dots \wedge \underline{x}_{\sigma(p)} = \text{sign}\sigma \ \underline{x}_1 \wedge \dots \wedge \underline{x}_p \tag{2.30}$$

Now suppose that $\underline{x}_1, \dots, \underline{x}_p$ are linearly dependent vectors. Then the skew symmetry of Λ^p implies that

$$\underline{x}_1 \wedge \dots \wedge \underline{x}_p = 0 \tag{2.31}$$

Conversely, p vectors which satisfy (2.31) are linearly dependent.

The results and definitions given above for general vector spaces will be specialised and discussed in more detail for the case of finite dimensional vector spaces.

Suppose that V is a vector space of dimension u over the field F. Then the p-th exterior power of V, $\Lambda^p V$ may always be defined ; $\Lambda^p V$ is a vector subspace of the p-th tensorial power of V. The pair $(\Lambda^p V, \Lambda^p)$ is uniquely defined up to an isomorphism. If $\underline{e}_i$, $i=1,\dots,\bar{n}$ is a basis of V then the products

$$\underline{e}_{i_1} \wedge \underline{e}_{i_2} \wedge \dots \wedge e_{i_p} \qquad i \leq i_1 \leq i_2 \leq i_p \leq n \tag{2.32}$$

span the vector space $\Lambda^p V$. There are $\binom{n}{p}$ choises of distinct indices $i_1, \dots, i_p$ from 1 to n, and they can be arranged uniquely in increasing order. It can be proved that the elements definec above, $\underline{e}_{i_1} \wedge \dots \wedge \underline{e}_{i_p}$ are linearly independent and thus form a basis for $\Lambda^p V$. Clearly then

$$\dim \Lambda^p V = \binom{n}{p}, \quad p=0,1,\dots,n \tag{2.33}$$

and $\Lambda^p V = 0$ for $p>n$.

An arbitrary vector of $\Lambda^p V$ is called a p-vector and an element of the form $\underline{x}_1 \wedge \dots \wedge \underline{x}_p$ where $\underline{x}_1, \dots, \underline{x}_p \in V$ is called decomposable. Every p-vector $\underline{u}$ of V can be uniquely represented in the form

$$\underline{u} = \sum_{<} a_{i_1 i_2 \dots i_p} \underline{e}_{i_1} \wedge \underline{e}_{i_2} \wedge \dots \wedge e_{i_p} \tag{2.34}$$

where the symbol < indicates that the indices $(i_1,\ldots,i_p)$ are ordered lexicographically $(1 \leq i_1 < i_2 < \ldots < i_p \leq n)$. The coefficients $a_{i_1 i_2 \ldots i_p}$ are called the coordinates of the p-vector $\underline{u}$ with respect to the basis $\{\underline{e}_i,\ i=1,\ldots,n\}$ of V.

2.3.1. Exterior powers of linear maps

Theorem (2.2) [Bir. -1]: Let V, U be finite dimensional vector spaces over a field F, and let $h: V \to U$ be a linear map. Then, there is a unique homomorphis, $\hat{h}: \Lambda V \to \Lambda U$ of the exterior algebras such that $\hat{h}(\underline{x}) = h(\underline{x})$ for any $\underline{x}$ in V. Notice that $\hat{h}$ maps $\Lambda^p V$ to $\Lambda^p U$ for all p.

The homomorphism $\hat{h}$ is a linear map. The above result simply means the following: If h is a linear map of a vector space V into a vector space U over F, then to $(\underline{x}_1,\ldots,\underline{x}_p) \in \overset{p}{\underset{1}{\times}} V$ we may correspond the element $h(\underline{x}_1) \wedge \ldots \wedge h(\underline{x}_p)$ of $\Lambda^p U$. This defines an alternating multilinear map ψ of $\overset{p}{\underset{1}{\times}} V$ into $\Lambda^p U$. By the definition of the exterior product there exists a unique linear map $\hat{h}$ of $\Lambda^p V$ into $\Lambda^p U$ ($\Lambda^p U$ is a vector space) such that

$$\hat{h}(\underline{x}_1,\ldots,\underline{x}_p) = h(\underline{x}_1) \wedge \ldots \wedge h(\underline{x}_p) \tag{2.35}$$

identically. The following commutative diagram describes the construction procedure for $\hat{h}$

$$\begin{tikzcd}
V \arrow[r, "h"] & U \\
\overset{p}{\underset{1}{\times}} V \arrow[r, "\Lambda^p"] \arrow[dr, "\psi"'] & \Lambda^p V \arrow[d, "\hat{h}=\Lambda^p h"] \\
& \Lambda^p U
\end{tikzcd}$$

Figure (2.2)

where we write $\Lambda^p h$ for $\hat{h}$ and we call it the p-th exterior power of linear map h. We have

$$\Lambda^p h(\underline{x}_1,\ldots,\underline{x}_p) = h(\underline{x}_1) \ \ldots \ h(\underline{x}_p) \tag{2.36}$$

Eqn. (2.36) defines the linear map $\Lambda^p h$ of $\Lambda^p V$ into $\Lambda^p U$. An important property of the map $\Lambda^p h$ are summarised below [Mar. -1].

□

Corollary (2.1): Let $f: V \to U$ and $g: U \to W$ be linear maps of finite dimensional vectör spaces over the field F. Set $h = g \circ f$. Then for the maps $\Lambda^p h$, $\Lambda^p f$, $\Lambda^p g$ we have

$$\Lambda^p(g \circ f) = \Lambda^p h = \Lambda^p h \circ \Lambda^p f \tag{2.37}$$

□

2.3.2. Representation theory of exterior powers of linear maps [Kar. -4]

Definitions and basic results [Ma. -1, Gre. -1]

Let V be an m-dimensional vector space over the field F and let $\Lambda^p V$, $p \leq m$ be the p-th exterior power of V. If $\{\underline{v}_i,\ i=1,\ldots,m\}$ is a basis of V, then $\Lambda^p V$ is spanned by the vectors of the basis $\{\underline{v}_\omega,\ \omega=(i_1,\ldots,i_p),\ 1 \leq i_1 < \ldots < i_p \leq m,\ \underline{v}_\omega = \underline{v}_{i_1} \wedge \underline{v}_{i_2} \wedge \ldots \wedge \underline{v}_{i_p}\}$. Every vector $\underline{v} \in \Lambda^p V$ pay be written as $\underline{v} = \Sigma_\omega a_\omega v_\omega$. Let r_v^p be the map of $\Lambda^p V$ into $F^{\binom{m}{p}}$ defined by

$$r_v^p(\underline{v}) = [\ldots, a_\omega, \ldots]^t \tag{2.38}$$

Then r_v^p is linear and it is called the representation map of $\Lambda^p V$ associated with the basis $\{\underline{v}_i,\ i=1,\ldots,m\}$. It can be seen that there is such a map associated to every basis of V. The image of $\Lambda^p V$ under this map is called the representation of $\Lambda^p V$ relative to the basis $\{\underline{v}_i,\ i=1,\ldots,m\}$ of V. The following result can be easily verified.

□

Proposition (2.7): The representations of the p-th exterior power of an m-dimensional vector space V over F, are isomorphisms of $\Lambda^p V$ onto $F^{\binom{m}{p}}$.

Let V, U be two vector spaces over the field F of dimensions m, n, respectively and let h be a linear map of V into U. The linear map h can be represented with respect to the bases $B_v = \{\underline{v}_i,\ i=1,\ldots,m\}$ and $B_u = \{\underline{u}_i,\ i=1,\ldots,n\}$

of $\mathcal{V}$ and $\mathcal{U}$ by a matrix H_u^v which is defined by the following commutative diagram

$$\begin{tikzcd}
\mathcal{V} \arrow[r, "h"] \arrow[d, "r_v^1"'] & \mathcal{U} \arrow[d, "r_u^1"] \\
F^m \arrow[r, "H_u^v"] & F^n
\end{tikzcd}$$

Figure (2.3)

where r_v^1, r_u^1 are the representation maps of $\mathcal{V}$ and $\mathcal{U}$ onto F^m and F^n respectively. Because $\mathcal{V}$, $\mathcal{U}$ are isomorphic to F^m, F^n respectively, F^m, F^n may be used to represent $\mathcal{V}$, $\mathcal{U}$ and the matrix H_u^v to represent the linear map h.

Let $\Lambda^p\mathcal{V}$, $\Lambda^p\mathcal{U}$ be the p-th exterior powers of $\mathcal{V}$, $\mathcal{U}$ respectively, where $p \leq \min\{m,n\}$. Then $h:\mathcal{V}\to\mathcal{U}$ implies the existence of a linear map $\Lambda^p h$ of $\Lambda^p\mathcal{V}$ into $\Lambda^p\mathcal{U}$. If we denote by r_v^p, r_u^p the representation maps of $\Lambda^p\mathcal{V}$, $\Lambda^p\mathcal{U}$ with respect to the bases $B_v=\{\underline{v}_i, i=1,\ldots,m\}$ and $B_u=\{\underline{u}_i, i=1,\ldots,n\}$ of $\mathcal{V}$, $\mathcal{U}$ respectively, then applying the representation result for linear maps, which has been discussed above, we have the following commutative diagram.

$$\begin{tikzcd}
\Lambda^p\mathcal{V} \arrow[d, "r_v^p"'] & \boxed{\underline{v}=\sum_\omega \alpha_\omega \underline{v}_\omega \Lambda} \arrow[r, "\Lambda^p h"] & \boxed{\underline{u}=\sum_\rho b_\rho \underline{u}_\rho \Lambda} & \Lambda^p\mathcal{U} \arrow[d, "r_u^p"] \\
F^{\binom{m}{p}} & \boxed{[\ldots,\alpha_\omega,\ldots]^t} \arrow[r, "\Lambda^p H_u^v"] & \boxed{[\ldots,b_\rho,\ldots]^t} & F^{\binom{n}{p}}
\end{tikzcd}$$

Figure (2.4)

and thus the matrix $\Lambda^p H_u^v$ is defined by the equation

$$\Lambda^p H_u^v \begin{bmatrix} \vdots \\ a_\omega \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ b_\rho \\ \vdots \end{bmatrix} \tag{2.39}$$

The description of $\Lambda^p H_u^v$ in terms of the H_u^v will be defined below and that will establish the links between the present subject and the compound matrix theory.

Let $B_v = \{\underline{v}_i,\ i=1,\dots,m\}$, $B_u = \{\underline{u}_i,\ i=1,\dots,n\}$ be bases of V and U respectively and let $\Lambda^p B_v = \{\underline{v}_\omega = \underline{v}_{i_1} \wedge \dots \wedge \underline{v}_{i_p},\ \omega = (i_1,\dots,i_p),\ 1 \le i_1 < \dots < i_p \le m\}$, $\Lambda^p B_u = \{\underline{u}_\rho = \underline{u}_{j_1} \wedge \dots \wedge \underline{u}_{j_p},\ \rho = (j_1,\dots j_p),\ 1 \le j_1 < \dots < j_p \le n\}$ be the induced bases of $\Lambda^p V$ and $\Lambda^p U$ respectively. Let

$$h(\underline{v}_i) = \sum_{j=1}^{n} c_{ij} \underline{u}_j, \quad i=1,\dots,m, \quad H_u^v = [c_{ij}] \tag{2.40}$$

Then for all basis vectors $\underline{v}_\omega \in \Lambda^p V$ we have

$$\Lambda^p h(\underline{v}_{i_1} \wedge \dots \wedge \underline{v}_{i_p}) = h(\underline{v}_{i_1}) \dots h(\underline{v}_{i_p}) =$$

$$\Big(\sum_{j=1}^{n} c_{i_1 j} \underline{u}_j\Big) \wedge \dots \wedge \Big(_{j=1} c_{i_p j} \underline{u}_j\Big) \tag{2.41}$$

But we know that

$$\underline{u}_{k_1} \wedge \dots \wedge \underline{u}_{k_p} = \operatorname{sign}\binom{j_1,\dots,j_p}{k_1,\dots,k_p} \underline{u}_{j_1} \dots \underline{u}_{j_p} \tag{2.42}$$

Also by the theory of determinants

$$H_{i_1,\dots,i_p}^{j_1,\dots,j_p} = \begin{vmatrix} c_{i_1 j_1} \dots c_{i_p j_1} \\ \\ c_{i_1 j_p} \dots c_{i_p j_p} \end{vmatrix} = \sum \operatorname{sign}\binom{j_1,\dots,j_p}{i_1,\dots,i_p} c_{i_1 j_1} \dots c_{i_p j_p} \tag{2.43}$$

Hence we have

$$\Lambda^p h(\underline{v}_{i_1} \wedge \dots \wedge \underline{v}_{i_p}) = \sum_{1 \le j_1 < \dots < j_p \le n} H_{i_1,\dots,i_p}^{j_1,\dots,j_p} \underline{u}_{j_1} \wedge \dots \wedge \underline{u}_{j_p} \tag{2.44}$$

Clearly, the quantities of the Eqn. (2.43) are the entries of the matrix, $\Lambda^p H_u^v$ which represents the linear map $\Lambda^p h:\Lambda^p V \to \Lambda^p U$ with respect to the bases $\Lambda^p B_v$, $\Lambda^p B_u$.

$\square$

2.3.3. Compound matrices and Grassman products [Mar. & Min. -1]

The results of the section 2.3.2. may be simplified by introducing some useful notation and definitions on the sequences of integers and on submatrices of a given matrix.

(i) Notation

(a) $Q_{p,n}$ denotes the set of strictly increasing sequences of p integers ($1 \leq p \leq n$) chosen from $1,\dots,n$, e.g. $Q_{2,3}=\{(1,2),\ (1,3),\ (2,3)\}$. Thus, the number of the sequences which belong to $Q_{p,n}$ is $\binom{n}{p}$.

If $\alpha, \beta \in Q_{p,n}$ we say that α precedes β ($\alpha<\beta$), if there exists an integer t ($1 \leq t \leq p$) for which $\alpha_1=\beta_1,\dots,\alpha_{t-1}=\beta_{t-1}$, $\alpha_t<\beta_t$, where α_i, β_i denote the elements of α, β respectively, e.g. in the set $Q_{3,8}$ $(3,5,8)<(4,5,6)$. This describes the lexicographic ordering of the elements of $Q_{p,n}$. The set of sequences $Q_{p,n}$ from now on will be assumed with its sequences lexicographically ordered and the elements of the ordered set $Q_{p,n}$ will be denoted by $Q_{p,n}(t)$, $t=1,\dots,\binom{n}{p}$ or simply by ω.

(b) If $c_1,\dots,c_n$ are elements of the field F and $\omega=(i_1,i_2,\dots,i_p)$ is a sequence in $Q_{p,n}$, $1 \leq p \leq n$, then the product $c_{i_1}c_{i_2}\dots c_{i_p}$ will be designated by c_ω.

(c) Suppose $A=[a_{ij}] \in M_{m,n}(F)$, where $M_{m,n}$ denotes the set of mxn matrices over the field F; let k, p be positive integers satisfying $1 \leq k \leq m$, $1 \leq p \leq n$ and let $\alpha=(i_1,\dots,i_k) \in Q_{k,m}$ and $\beta=(j_1,\dots,j_p) \in Q_{p,n}$. Then $A[\alpha|\beta]$ $M_{k,p}(F)$ denotes the submatrix of A which contains the rows $i_1,\dots,i_k$ and the columns $j_1,\dots,j_p$. We use the notation $A(\alpha|\beta]$ to designate the submatrix of A which excludes rows $i_1,\ldots,i_k$ and includes columns $j_1,\ldots,j_p$. The submatrices $A[\alpha|\beta)$ and $A(\alpha|\beta)$ can be defined similarly.

□

(ii) Compound matrices

Let $A \in M_{m,n}(F)$ and $1 \leq p \leq \min\{m,n\}$, then the p-th compound matrix or p-th adjugate of A is the $\binom{m}{p} \times \binom{n}{p}$ matrix whose entries are $\det\{A[\alpha|\beta]\}$, $\alpha \in Q_{p,m}$, $\beta \in Q_{p,n}$ arranged lexicographically in α and β. This matrix will be designated by $C_p(A)$. For example if $A \in M_3(F)$ and p=2, then $Q_{2,3}=\{(1,2), (1,3), (2,3)\}$ and

$$C_2(A) = \begin{bmatrix} \det\{A[(1,2)|(1,2)]\} & \det\{A[(1,2)|(1,3)]\} & \det\{A[(1,2)|(2,3)]\} \\ \det\{A[(1,3)|(1,2)]\} & \det\{A[(1,3)|(1,3)]\} & \det\{A[(1,3)|(2,3)]\} \\ \det\{A[(2,3)|(1,2)]\} & \det\{A[(2,3)|(1,3)]\} & \det\{A[(2,3)|(2,3)]\} \end{bmatrix} \quad (2.45)$$

□

Properties of compound matrices

(a) If $A \in M_n(F)$, $1 \leq p \leq n$ and also A is non-singular

(i) $[C_p(A)]^{-1} = C_p(A^{-1})$ (2.46)

(ii) $C_p(A) = [C_p(A)]^*$, where A* is the conjugate transpose of $A(F=\mathbb{C})$ (2.47)

(iii) $C_p(A^t) = [C_p(A)]^t$, where A^t is the transpose of A (2.48)

(iv) $C_p(\bar{A}) = \overline{C_p(A)}$, where $\bar{A}$ is the conjugate of $A(F=\mathbb{C})$ (2.49)

(v) $C_p(kA) = k^p C_p(A)$, for any $k \in F$ (2.50)

(vi) $C_p(I_n) = I_{\binom{n}{p}}$ (2.51)

(vii) $\det\{C_p(A)\} = \{\det A\}^{\binom{n-1}{p-1}}$ Sylvester-Franke Theorem (2.52)

(b) If $A \in M_{m,n}(F)$ and $B \in M_{n,k}(F)$ and $1 \leq p \leq \min\{m,n,k\}$, then

$$C_p(AB) = C_p(A)C_p(B) \quad \text{Binet-Cauchy Theorem} \quad (2.53)$$

(c) If $A \in M_{p,n}(F)$ and the p rows of A are denoted by $\underline{a}_1^{-t},\dots,\underline{a}_p^{-t}$ in succession ($1 \le p \le n$), then $C_p(A)$ is an $\binom{n}{p}$-tuple and it is called the Grassmann product or skew symmetric product of the vectors $\underline{a}_1^{-t},\dots,\underline{a}_p^{-t}$ for reasons which will become apparent later on. The usual notation for this $\binom{n}{p}$-tuple of subdeterminants of A is $\underline{a}_1^{-t}\wedge\dots\wedge\underline{a}_p^{-t}$ and it denotes a row vector. The Grassmann product of the columns of a matrix $A \in M_{n,p}(F)$ ($1 \le p \le n$) may be defined in a similar manner; the product in this case, however, will be an $\binom{n}{p}$-column vector. If $\underline{a}_1,\dots,\underline{a}_p$ are the columns of A, in this case, then this $\binom{n}{p}$-tuple of subdeterminants of A will be denoted by $\underline{a}_1\wedge\dots\wedge\underline{a}_p$. By the properties of determinants, if $\sigma \in S_p$ (where S_p denotes the totality of permutations of $(1,\dots,p)$), then

$$\underline{a}_{\sigma(1)}\wedge\dots\wedge\underline{a}_{\sigma(p)} = \operatorname{sign}\sigma\ \underline{a}_1\wedge\dots\wedge\underline{a}_p \tag{2.54}$$

If $B \in M_n(F)$, $A \in M_{n,p}(F)$, then by the Binet-Cauchy theorem it follows that

$$C_p(B)\underline{a}_1\wedge\dots\wedge\underline{a}_p = B\underline{a}_1\wedge\dots\wedge B\underline{a}_p \tag{2.55}$$

Grassmann products suitably deployed may greatly reduce the complexity of the expressions in compound matrices. Thus, let $A \in M_{m,n}(F)$ and $1 \le p \le \min\{m,n\}$. The matrix A may be written in terms of its rows or columns respectively as

$$A = \begin{bmatrix} \underline{\bar{a}}_1^t \\ \vdots \\ \underline{\bar{a}}_m^t \end{bmatrix} \quad \text{or } A = [\underline{a}_1,\dots,\underline{a}_n] \tag{2.56}$$

Let $\omega=(i_1,\dots,i_p) \in Q_{p,m}$ and $\phi=(j_1,\dots,j_p) \in Q_{p,n}$ and let us denote by $\underline{\bar{a}}_\omega^t$ the Grassmann product $\underline{\bar{a}}_{i_1}^t \dots \underline{\bar{a}}_{i_p}^t$ and by $\underline{a}_\phi$ the Grassmann product $\underline{a}_{j_1} \dots \underline{a}_{j_p}$. The p-th compound matrix of A may then be expressed in either of the following forms

$$C_p(A) = \begin{bmatrix} \vdots \\ \bar{a}^t \\ \vdots \end{bmatrix}, \quad \omega \in Q_{p,m} \quad \text{or} \quad C_p(A) = [\dots,\underline{a}_\phi,\dots] \tag{2.57}$$

which will be referred to as row, columns representations of $C_p(A)$ respectively.

□

2.3.4. Compound matrices and exterior algebra [Kar. -4], [Gian. -1]

We may now return to the Eqn. (2.44). We first note that the matrix H_u^v defined by (2.40) or by

$$[h(\underline{v}_1), h(\underline{v}_2), \dots, h(\underline{v}_m)] = [\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n] \begin{bmatrix} c_{11} c_{21} \dots c_{m1} \\ c_{12} c_{22} \dots c_{m2} \\ \vdots \quad \vdots \qquad \vdots \\ c_{1n} c_{2n} \dots c_{mn} \end{bmatrix} \tag{2.58}$$

is the matrix representation of $h: V \to U$ with respect to the bases B_v, B_u of V, U respectively. Note that

$$H_{i_1 \dots i_p}^{j_1 \dots j_p}$$

is the p-th order minor of H_u^v that lies on the $j_1, j_2, \dots, j_p$ rows and $i_1, i_2, \dots, i_p$ columns. If we define $\underline{h}_\omega \wedge = \underline{h}_{i_1} \wedge \underline{h}_{i_2} \wedge \dots \wedge \underline{h}_{i_p}$, where $\underline{h}_{i_1}, \underline{h}_{i_2}, \dots, \underline{h}_{i_p}$ are the columns of H_u^v that correspond to the indices $(i_1, i_2, \dots, i_p) \in Q_{p,m}$, then Eqn. (2.44) may be written as

$$\Lambda^p h(\underline{v}_\omega \wedge) = [\dots, \underline{u}_\rho \wedge, \dots] \, \underline{h}_\omega \wedge, \quad \rho \in Q_{p,n} \tag{2.59}$$

Given that the relationship holds for all $\omega \in Q_{p,m}$, $\{\underline{v}_\omega \wedge, \omega \in Q_{p,m}\}$ is a basis of $\Lambda^p V$ and $\{\underline{u}_\rho \wedge, \rho \in Q_{p,n}\}$ is a basis of $\Lambda^p U$, we may write

$$[\dots, \Lambda^p h(\underline{v}_\omega \wedge), \dots] = [\dots, \underline{u}_\rho \wedge, \dots] \, [\dots, \underline{h}_\omega \wedge, \dots] = \Lambda^p B_u \Lambda^p H_u^v = \Lambda^p B_u C_p(H_u^v) \tag{2.60}$$

where $\Lambda^p H_u^v = C_p(H_u^v)$ is the matrix representation of $\Lambda^p h$ with respect to the bases $\Lambda^p B_v$, $\Lambda^p B_u$, and it is defined that by the p-th compound matrix of H_u^v. These considerations lead to the following result.

□

<u>Theorem (2.3)</u>: Let $\mathcal{V}$, $\mathcal{U}$ be two vector spaces over $\mathcal{F}$, with $\dim\mathcal{V}=m$, $\dim\mathcal{U}=n$ and let $h:\mathcal{V}\to\mathcal{U}$ be a linear map of $\mathcal{V}$ into $\mathcal{U}$. Let $B_v=\{\underline{v}_i,\ i=1,\dots,m\}$, $B_u=\{\underline{u}_j,\ j=1,\dots,n\}$ be bases of $\mathcal{V}$, $\mathcal{U}$ respectively and let H_u^v be the matrix representation of h with repsect to the bases B_v, B_u. If $\Lambda^p h:\Lambda^p\mathcal{V}\to\Lambda^p\mathcal{U}$, $1\leq p\leq\min\{m,n\}$, is the p-th exterior power of h, then $\Lambda^p h$ may be represented with respect to the induced bases $\Lambda^p B_v=\{\underline{v}_\omega\wedge,\ \omega\in Q_{p,m}\}$, $\Lambda^p B_u=\{\underline{u}_\rho\wedge,\ \rho\in Q_{p,n}\}$ of $\Lambda^p\mathcal{V}$, $\Lambda^p\mathcal{U}$ repsectively, by the matrix $\Lambda^p H_u^v=C_p(H_u^v)$ where $C_p(H_u^v)$ is the p-th compound matrix of H_u^v.

The above result can be represented by the following commutative diagram

$$\begin{tikzcd}
B_v,\mathcal{V} \arrow[r, "h"] \arrow[dd, "r_v^1"'] & \mathcal{U},B_u \arrow[dd, "r_u^1"] & & \Lambda^p B_v,\Lambda^p\mathcal{V} \arrow[r, "\Lambda^p h"] \arrow[dd, "r_v^p"'] & \Lambda^p\mathcal{U},\Lambda^p B_u \arrow[dd, "r_u^p"] \\
 & & \Longrightarrow & & \\
\mathcal{F}^m \arrow[r, "H_u^v"] & \mathcal{F}^n & & \mathcal{F}^{\binom{m}{p}} \arrow[r, "C_p(H_u^v)"] & \mathcal{F}^{\binom{n}{p}}
\end{tikzcd}$$

Figure (2.5)

It has been shown that the pairs of the vectors spaces $(\Lambda^p\mathcal{V},\mathcal{F}^{\binom{m}{p}})$ and $(\Lambda^p\mathcal{U},\mathcal{F}^{\binom{n}{p}})$ are isomorphic. In fact, every basis B_v of $\mathcal{V}$ and B_u of $\mathcal{U}$ induces a decomposable basis for $\Lambda^p\mathcal{V}$, $\Lambda^p\mathcal{U}$ and the corresponding representation maps r_v^p, r_u^p define isomorphisms between $\Lambda^p\mathcal{V}$, $\mathcal{F}^{\binom{m}{p}}$ and $\Lambda^p\mathcal{U}$, $\mathcal{F}^{\binom{n}{p}}$. The linear map $\Lambda^p H_u^v=C_p(H_u^v):\mathcal{F}^{\binom{m}{p}}\to\mathcal{F}^{\binom{n}{p}}$ is induced by the map $H_u^v:\mathcal{F}^m\to\mathcal{F}^n$ and it is a representation of the linear map $\Lambda^p h:\Lambda^p\mathcal{V}\to\Lambda^p\mathcal{U}$. Thus, it is clear that, as any pair of vector spaces $\mathcal{V}$, $\mathcal{U}$ of finite dimension and their linear map h can be discussed by means of m-tuples, n-tuples and matrices, their p-th exterior powers $\Lambda^p\mathcal{V}$, $\Lambda^p\mathcal{U}$ and their linear map $\Lambda^p h$ may be discussed in terms of $\binom{m}{p}$-tuples, $\binom{n}{p}$-tuples and compound matrices.

Let $L(\mathcal{V},\mathcal{U})$ be the vector space of linear maps of the m-dimensional vector space $\mathcal{V}$ into the n-dimensional vector space $\mathcal{U}$, both spaces being defined over the field F. Let $M_{n,m}$ be the set of nxm matrices over F; then each matrix representation of an element of $L(V,U)$ is an element of $M_{n,m}$. It is known that the map which is defined by matrix representation is an isomorphism of $L(V,U)$ onto $M_{n,m}$ [Bir. -1]. In diagramatic terms we may represent the above discussion as follows.

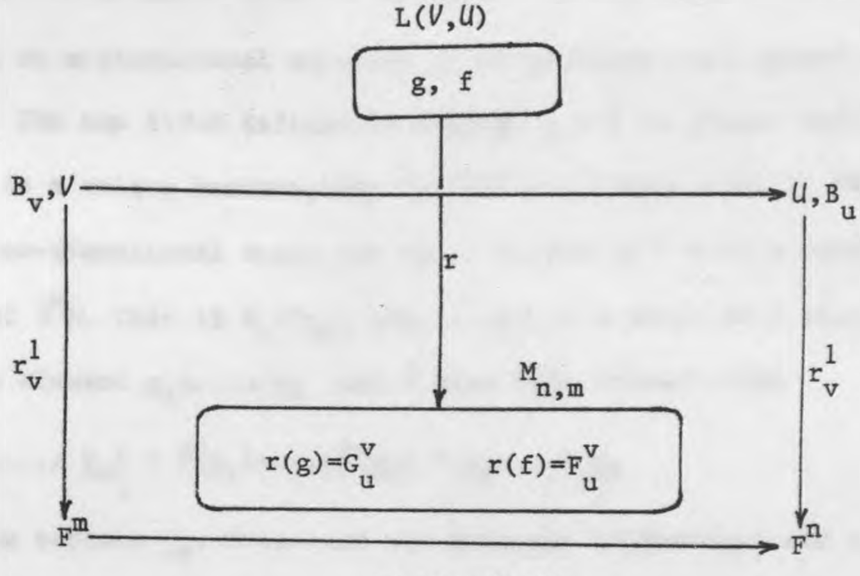

Figure (2.6)

□

Proposition (2.8) [Kar. -4]: Let V, U be two vector spaces over the field F, $\dim V=m$, $\dim U=n$ and let B_v, B_u be bases of V, U repsectively.

(i) The set $L_p(\Lambda^p V,\Lambda^p U)$ of linear maps of $\Lambda^p V$ into $\Lambda^p U$ is a vector space.

(ii) The map r_p that associates every map $\Lambda^p h$ with its matrix representation $C_p(H_u^v)$ is an isomorphism of $L_p(\Lambda^p V,\Lambda^p U)$ onto the set of matrices $M_{\binom{n}{p},\binom{m}{p}}$.

In diagramatic terms we may represent this result as follows:

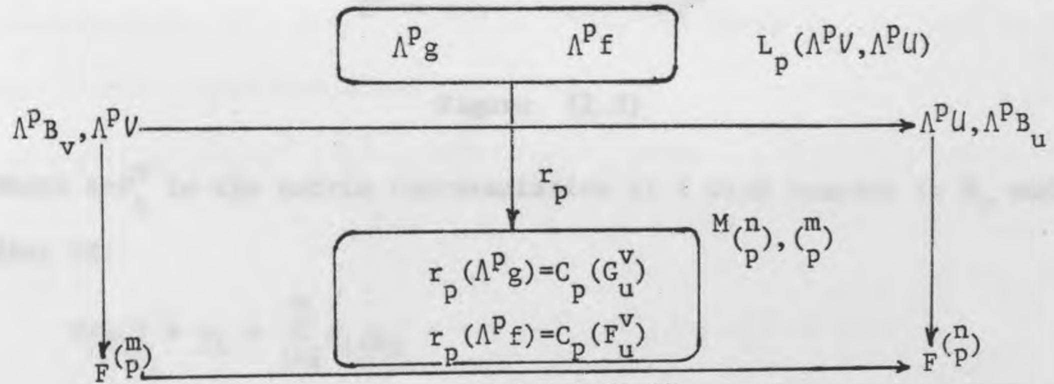

Figure (2.7)

Using the above results, it can be readily shown [Kar. -4] that the Binet-Cauchy Theorem, in its compound matrix form expressed the composition law of the exterior powers of linear maps, when matrix representations are considered.

□

2.3.5. Plücker coordinates and decomposability [Mar. -1],[Gre. -1],[Hod. & Bed. -1]

Let V be an m-dimensional subspace of an n-dimensional vector space U over a field F. The map $f:V\to U$ defined by $f(\underline{x})=\underline{x}$, $\underline{x}\in V$ is linear and by theorem 2.2 there is a unique homomorphism $\hat{f}:\Lambda V\to\Lambda U$ associated with f. Since $\dim V=m$, $\Lambda^m V$ is a one-dimensional space and it is mapped by $\hat{f}$ onto a one-dimensional subspace of $\Lambda^m U$. This if $B_v=\{\underline{v}_i, i=1,\ldots,m\}$ is a basis of V then $\Lambda^m V$ is spanned by the element $\underline{v}_1\wedge\ldots\wedge\underline{v}_m$ and $\hat{f}$ maps this element onto

$$\hat{f}(\underline{v}_1\wedge\ldots\wedge\underline{v}_m) = \hat{f}(\underline{v}_1)\wedge\ldots\wedge\hat{f}(\underline{v}_m) = \underline{v}_1\wedge\ldots\wedge\underline{v}_m \tag{2.61}$$

in $\Lambda^m U$. The vectors $\underline{v}_i$, i=1,...,m are linearly independent and so $\underline{v}_1\wedge\ldots\wedge\underline{v}_m$ is a non-zero element of $\Lambda^m U$. In fact the injection map $f:V\to U$ defined by $f(\underline{x})=\underline{x}$, $\underline{x}\in V$ induces an injection map $\Lambda^m f:\Lambda^m V\to\Lambda^m U$ defined by $\Lambda^m f(\underline{x}\wedge)=\underline{x}\wedge$, $\underline{x}\wedge\in\Lambda^m V$. The vector $\underline{v}_1\wedge\ldots\wedge\underline{v}_m$ spans a one-dimensional subspace of $\Lambda^m U$ which depends only on V. Now let $B_u=\{\underline{u}_j, j=1,\ldots,n\}$ be a basis of U, then using matrix representations we have the following commutative diagram

$$\begin{tikzcd}
B_v,V \arrow[r, "f"] \arrow[d, "r_v^1"'] & U,B_u \arrow[d, "r_u^1"] \\
F^m \arrow[r, "A=F_u^v"] & F^n
\end{tikzcd}$$

Figure (2.8)

where $A=F_u^v$ is the matrix representation of f with respect to B_v and B_u. In fact if

$$f(\underline{v}_i) = \underline{v}_i = \sum_{j=1}^{n} a_{ij}\underline{u}_j \tag{2.62}$$

The column span of A is a subspace of $\mathcal{F}^n$ and it is the representation of $f(\mathcal{V})$ with respect to the bases B_u, B_v. The representation of $\underline{v}_i\wedge..\wedge\underline{v}_m$ with respect to the bases $\Lambda^m B_u$, $\Lambda^m B_v$ is defined by the commutative diagram

$$\begin{tikzcd}
\Lambda^m B_v, \Lambda^m \mathcal{V} \arrow[r, "\Lambda^m f"] \arrow[d, "r_v^m"'] & \Lambda^m \mathcal{U}, \Lambda^m B_u \arrow[d, "r_u^m"] \\
F = F^{\binom{m}{m}} \arrow[r, "C_m(F_u^v)=C_m(A)"] & F^{\binom{n}{m}}
\end{tikzcd}$$

Figure (2.9)

Thus

$$\underline{v}_1\wedge\cdots\wedge\underline{v}_m = \Sigma a_\omega \underline{u}_\omega \wedge = [\ldots,\underline{u}_\omega\wedge,\ldots]\begin{bmatrix}\vdots\\ a_\omega\\ \vdots\end{bmatrix}, \quad \omega=(i_1,\ldots,i_m)\in Q_{m,m} \tag{2.63}$$

and hence the matrix

$$C_m(A) = C_m(F_u^v) = \underline{a}_1\wedge\ldots\wedge\underline{a}_m = \begin{bmatrix}\vdots\\ a_\omega\\ \vdots\end{bmatrix} \in \mathbb{R}^{\binom{n}{m}\mathrm{x}1} \tag{2.64}$$

is the "matrix" representation of $\Lambda^m f$ with respect to $\Lambda^m B_u$, $\Lambda^m B_v$. The $\binom{n}{m}$-tuple $(\ldots,a_\omega,\ldots)$ are the coordinates of the one-dimensional subspace $\Lambda^m f(\Lambda^m \mathcal{V})$ with respect to the two given bases. If $B_v=\{\underline{v}_i,\ i=1,\ldots,m\}$ and $B_{v'}=\{\underline{v}_i',\ i=1,\ldots,m\}$ are two bases of $\mathcal{V}$ and B_u is a fixed basis of $\mathcal{U}$, then the matrix representations of f with respect to those two bases B_v, $B_{v'}$ are related by the coordinate transformation $Q_{v'}^v$,

$$F_u^{v'} = F_u^v Q_{v'}^v,\ F_u^{v'},\ F_u^v \in \mathcal{F}^{n\mathrm{x}m},\quad Q_{v'}^v \in \mathcal{F}^{m\mathrm{x}m} \tag{2.65}$$

and thus

$$C_m(F_u^{u'}) = C_m(F_u^v)C_m(Q_{v'}^v) = C_m(F_u^v)q,\quad q=\det\{Q_u^v\}\in \mathcal{F}-\{0\} \tag{2.66}$$

The two vectors $\underline{t}=\underline{v}_1\wedge\ldots\wedge\underline{v}_m$, $\underline{t}'=\underline{v}_1'\wedge\ldots\wedge\underline{v}_m'$ are related by

$$\underline{t}' = [\ldots,\underline{u}_\omega\wedge,\ldots]\begin{bmatrix}\vdots\\ a_\omega'\\ \vdots\end{bmatrix} = [\ldots,\underline{u}_\omega\wedge,\ldots]\begin{bmatrix}\vdots\\ a_\omega\\ \vdots\end{bmatrix} q = q\underline{t} \tag{2.67}$$

or

$$a_\omega' = a\, a_\omega' \quad \omega \in Q_{m,n} \tag{2.68}$$

□

Definition (2.18): The scalars a of Eqn. (2.64) are called *Plücker coordinates of the subspace V* relative to the bases B_v of V and B_u of U.

Eqn. (2.68) shows that any two sets of Plücker coordinates of V, which correspond to two different bases of V, with respect to the fixed basis B_u of U differ by a non-zero scalar factor. Hence the ratios of a_ω's are the same as the corresponding ratios of a_ω''s ($a_{\omega_1}=qa_{\omega_1}'$, $a_{\omega_2}=qa_{\omega_2}'$ and so $a_{\omega_1}|a_{\omega_2}=a_{\omega_1}'|a_{\omega_2}'$). Therefore, the ratios are uniquely determined by V. Sometimes, the ratios of the a_ω, rather that the a_ω themselves, are called the Plücker coordinates of V.

Consider now the vector space F^{p+1} of (p+1)-tuples $\underline{x}=(x_o,x_1,\ldots,x_p)$, $x_i \in F$. Let us call two such vectors $\underline{x}$ and $\underline{y}$ equivalent if they are both non-zero and if $\underline{x}=q\underline{y}$ for some $q \in F-\{0\}$. This equivalence relation splits the non-zero vectors in F^{p+1} into equivalence classes, and clearly each equivalence class consists of all non-zero elements in a one-dimensional subspace of F^{p+1}. Thus the equivalence classes are in one-to-one correspondence with the lines through the origin of F^{p+1}.

□

Definition (2.19): The set of all equivalence classes of non-zero vectors in F^{p+1} as defined above, is called the *projective space of dimension p over F*, denoted by $\mathbb{P}^p(F)$. Each equivalence class defines a point of this projective space. If Q is any point in $\mathbb{R}^p(\mathcal{F})$ and if $\underline{x}=(x_o,\dots,x_p)$ is any vector of the equivalence class which defines Q, then the $x_i^{'}$s are called <u>homogeneous coordinates</u> of Q.

If we set $p=\binom{n}{m}-1=\dim\Lambda^m\mathcal{U}-1$, then we can easily see that the Plücker coordinates of $\mathcal{V}$, enumerated in lexicographic order, may be considered as the homogeneous coordinates of a point in $\mathbb{P}^p(\mathcal{F})$. However, every point in $\mathbb{P}^p(\mathcal{F})$ does not represent an m-dimensional subspace of $\mathcal{U}$. Elements of $\Lambda^m\mathcal{U}$ of the type $q\underline{v}_1\Lambda\dots\Lambda\underline{v}_m$ where $\underline{v}_1,\dots,\underline{v}_m$ are linearly independent vectors of $\mathcal{V}$ and $q \in \mathcal{F}-\{0\}$ are called <u>simple</u> or <u>decomposable m-vectors</u>. Decomposable multivectors uniquely define m-dimensional subspaces of $\mathcal{U}$ as it is shown below.

□

<u>Proposition (2.9)</u>: Let $\mathcal{U}$ be an n-dimensional vector space over $\mathcal{F}$ and let $\underline{y} = \underline{y}_1\Lambda\dots\Lambda\underline{y}_m$, $\underline{z} = \underline{z}_1\Lambda\dots\Lambda\underline{z}_m$ be two decomposable non-zero elements of $\Lambda^m\mathcal{U}$ and let us denote by $\mathcal{V}_y=\text{span}\{\underline{y}_1,\dots,\underline{y}_m\}$ and $\mathcal{V}_z=\text{span}\{\underline{z}_1,\dots,\underline{z}_m\}$ the subspaces of $\mathcal{U}$ defines by $\underline{y}\Lambda$ and $\underline{z}\Lambda$ respectively. Necessary and sufficients condition for $\mathcal{V}_y=\mathcal{V}_z$ is

$$\underline{y}\Lambda = \underline{y}_1\Lambda \quad \Lambda\underline{y}_m = q\underline{z}_1\Lambda\dots\Lambda\underline{z}_m = q\underline{z}\Lambda \ , \quad q \in \mathcal{F}-\{0\} \tag{2.69}$$

□

<u>Definition (2.20)</u>: Let $\mathcal{U}$ be a vector space over a field $\mathcal{F}$ with $\dim\mathcal{U}=n$. The <u>Grassmannian $G(m,\mathcal{U})$</u> is defined as the set of m-dimensional subspaces $\mathcal{V}$ of $\mathcal{U}$; $G(m,\mathcal{U})$ actually admits the structure of an analytic manifold which is known as the <u>Grassmann manifold</u>.

It is clear that the mapping $f:G(m,\mathcal{U})\to G(1,\Lambda^m\mathcal{U})$ expresses a natural injective correspondence between $G(m,\mathcal{U})$ and the set of one-dimensional subspaces of $\Lambda^m\mathcal{U}$ (i.e. $G(1,\Lambda^m\mathcal{U})$).

□

<u>Example (2.1)</u>: To demonstrate that every vector of $\Lambda^m\mathcal{U}$ does not define an m-dimensional subspace $\mathcal{V}$ of $\mathcal{U}$, we consider the simple case of $\dim\mathcal{U}=n=4$, $\dim\mathcal{V}=m=2$. Let $B_v=\{\underline{v}_1,\underline{v}_2\}$ be a basis of $\mathcal{V}$. Then, as ir is well known, we can

extend B_v to $B_u=\{\underline{v}_1,\underline{v}_2,\underline{v}_3,\underline{v}_4\}$ which is a basis of U. Thus the induced basis of $\Lambda^2 U$ is $\Lambda^2 B_u=\{\underline{v}_1\wedge\underline{v}_2,\underline{v}_1\wedge\underline{v}_3,\underline{v}_1\wedge\underline{v}_4,\underline{v}_2\wedge\underline{v}_3,\underline{v}_2\wedge\underline{v}_4,\underline{v}_3\wedge\underline{v}_4\}$ and so

$$\underline{v} = \underline{v}_1\wedge\underline{v}_2 = \sum_{1\le i<j\le 4} c_{ij}\underline{v}_i\ \underline{v}_j \tag{2.70}$$

where c_{ij}, $1\le i<j\le 4$ is a set of Plücker coordinates of V. It can be shown that $\underline{V}$ is decomposable [Mar. -1] if and only if

$$c_{12}c_{34} - c_{13}c_{24} + c_{14}c_{23} = 0 \tag{2.71}$$

Clearly, the above condition is a necessary condition for the general 6-tuple $(c_{12},c_{13},c_{14},c_{23},c_{24},c_{34})$ to be the Plücker coordinates of a 2-dimensional subspace V, or in other words to be the coordinates of a decomposable vector. This condition may also be proved sifficient.

Such a condition is knows as a Quadratic Plücker Relationship. This Quadratic Plücker Relationship defines a hypersurface in the 5-dimensional projective space $\mathbb{P}^5(F)$, which is known as the Grassmann variety of this projective space.

In general, those points of $\mathbb{P}^p(F)$, $p=\binom{n}{m}-1$ which correspond to m-dimensional subspaces V of the n-dimensional vector space U must satisfy a set of quadratic relationships of the type (2.71). It will be seen that this set of relationship defines an algebraic variety of the projective space $\mathbb{P}^p$ (F) which is known as the Grassmann variety of $\mathbb{P}^p(F)$.

By associating to every $V \in G(m,U)$ its Plücker coordinates $(\dots,a_\omega,\dots)$, $\omega \in Q_{m,n}$ the map $g:G(m,U)\to\mathbb{P}^p(F)$ is defined, and this is known as the Plücker embedding of $G(m,U)$ in the projective space $\mathbb{P}^p(F)$. The Plücker image of $G(m,U)$ in $\mathbb{P}^p(F)$ the above defined Grassmann variety if $\mathbb{P}^p(F)$ [Hod. -1].

Finally if V is any m-dimensional subspace of the n-dimensional vector space U, then any non-zero decomposable multivector $\underline{v}_1\wedge\dots\wedge\underline{v}_m$ with $\underline{v}_i \in V$, i=1,2. ...,m is called a Grassmann representative of V. We have already seen that all the Grassmann representatives differ only by non-zero scalar factors so that we shall denote any one of them simply by $\underline{g}(V)$.

□

2.4. Segre characteristic of $A \in \mathbb{R}^{nxn}$, Ferrer's diagram, Jordan form [Tur. & Ait. -1]

2.4.1. Generalized Nullspace and Range of an operator

Let $T:V \to V$ be a linear operator on an n-dimensional vector space V and let A the nxn matrix representation of T.

If we define: $N_i = N_r(T^i)$, i=1,2,... then this sequence of subspaces is non-decreasing, because $N_i \subseteq N_{i+1}$, $\forall$ i $\in \mathbb{N}$ and because V is finite dimensional there exist the supremal element, i.e., $\exists q \in \mathbb{N}$ such that $N_q = N_{q+1} = \ldots$

Definition (2.21) [Dorny -1]: The Generalized Nullspace $N_g(T) \triangleq N_q$ of T is the largest subspace of V annihilated by powers of T, i.e.

$$N_g(T) \triangleq N_q = N_r(T^q) \tag{2.72}$$

The power q of T required for maximum annihilation is called the index of annihilation of T.

The generalised range $R_g(T)$ of the operator T is defined by:

$$R_g(T) \triangleq \text{range}(T^q) = R(T^q) \tag{2.73}$$

then both $N_g(T)$ and $R_g(T)$ are T-invariant subspaces and $V = N_g(T) \oplus R_g(T)$

□

2.4.2. Generalised Eigenspaces

Let A be an nxn matrix with the characteristic polynomial, $\varphi(\lambda) = \det(\lambda I_n - A) = (\lambda-\lambda_1)^{\tau_1}(\lambda-\lambda_2)^{\tau_2}\ldots(\lambda-\lambda_\rho)^{\tau_\rho}$, where $\lambda_i \neq \lambda_j$ and $\tau_1+\tau_2+\ldots+\tau_\rho = n$.

Define, $$U_i \triangleq N_g(A-\lambda_i I) = N_r\{(A-\lambda_i I)^{q_i}\} \triangleq N_g(A,\lambda_i) \tag{2.74}$$

where q_i is the index of annihilation of $A-\lambda_i I$, then U_i will be called the generalised eigenspace of A at λ_i and of course is the largest subspace of $\mathbb{R}^n$ (or $\mathbb{C}^n$) annihilated by powers of $A-\lambda_i I$.

Properties of generalised eigenspaces

(i) $N_g(A,\lambda_i)\neq 0$ iff $\lambda_i \in \sigma(A)$

(ii) If $U_i=N_g(A,\lambda_i)$, $\lambda_i \in \sigma(A)$, them $\dim U_i=\tau_i$=algebraic multiplicity of λ_i.

(iii) $\{U_i=N_g(A,\lambda_i),\ \lambda_i \in \sigma(A),\ i=\underset{\sim}{\rho}\}$ are linearly independent $\Leftrightarrow U_i \cap (\sum_{j,j\neq i} U_j)=$

$=\{0\}$ and $\mathbb{R}^n$ (or $\mathbb{C}^n$)$=\oplus \sum_{i=1}^{\rho} U_i$.

□

2.4.3. Generalised Eigenvectors and chains of Generalised Eigenvectors

Definition (2.22): $\underline{u}_k$ is called a generalised eigenvector of A of rank k for λ_i iff $(A-\lambda_i I)^k \underline{u}_k=0$ and $(A-\lambda_i I)^{k-1}\underline{u}_k\neq 0$.

Thus eigenvectors of A for λ_i are generalised eigenvectors of A for λ_i of rank 1.

Theorem (2.4): Let $U_i=N_g(A-\lambda_i I)$, $\lambda_i \in \sigma(A)$, then generalised eigenvectors of A corresponding to λ_i but of differing ranks are linearly independent.

□

Theorem (2.5): Let $A \in \mathbb{R}^{n\times n}$, $\lambda_i \in \sigma(A)$ and $U_i=N_g(A-\lambda_i I)$. if $\underline{u}_k$ is a generalised eigenvector of A for λ_i, of rank k, k>1 then $(A-\lambda_i I)\ \underline{u}_k$ is a generalised eigenvector of A for λ_i and of rank k-1.

□

Hence a generalised eigenvector $\underline{u}_k$ of rank k, k>1 for λ_i can be used to generate a chain of k linearly independent generalised eigenvectors of ranks k,k-1,...,1. They are the vectors

$$\underline{u}_k,\ (A-\lambda_i I)\underline{u}_k,\ (A-\lambda_i I)^2\underline{u}_k,\dots,\ (A-\lambda_i I)^{k-1}\underline{u}_k \tag{2.74}$$

The final vector in the chain, namely $(A-\lambda_i I)^{k-1}\underline{u}_k$ will be an eigenvector since it is a non zero vector that satisfies the identity

$$(A-\lambda_i I)(A-\lambda_i I)^{k-1}\underline{u}_k = 0$$

A chain is said to be a maximal chain if its elements cannot be considered a proper subset of another chain.

Maximal chains must exist since the vectors that make up a chain are linearly independent and the generalised eigenspace in which they lie is of finite dimension.

Note that the index of annihilation q_i is the largest rank of any generalised eigenvector in $\mathcal{U}_i$, also there may be more than one maximal chains in $\mathcal{U}_i$ consisting of q_i vectors and there may be maximal chains in $\mathcal{U}_i$ consisting of fewer than q_i vectors.

It can be seen that if q_i is the index of annihilation then $\mathcal{U}_i$ is kernel of $(A-\lambda_i I)^{q_i}$ (Every element of this kernel is a generalised eigenvector corresponding to λ_i and every generalised eigenvector is an element of this kernel).

<u>Theorem (2.6)</u>: Assume $\underline{u}_k^{\alpha}$, $\underline{u}_{\ell}^{\beta}$ generalised eigenvectors of A for λ of ranks k, ℓ and the corresponding chains

$$C_k^{\alpha} = \{\underline{u}_i^{\alpha} : \underline{u}_i^{\alpha} = (A-\lambda I)^{k-i}\underline{u}_k^{\alpha}, \quad i=1,2,\dots,k\} \tag{2.75}$$

$$C_{\ell}^{\beta} = \{\underline{u}_j^{\beta} : \underline{u}_j^{\beta} = (A-\lambda I)^{\ell-j}\underline{u}_{\ell}^{\beta}, \quad j=1,2,\dots,\ell\} \tag{2.76}$$

If $\underline{u}_1^{\alpha}$, $\underline{u}_1^{\beta}$, the true eigenvectors are linearly independent, then C_k^{α}, C_{ℓ}^{β} are linearly independent.

□

Note, that the number of independent chains d of $\mathcal{U}_i$ is d=n-rank$(A-\lambda_i I)$= number of independent eigenvectors.

In order to find the number of elements in each of the independent chains we can use the Weyr characteristic and the Ferrer's diagram [Tur. & Ait. -1].

Let $A \in \mathbb{R}^{n\times n}$, $\varphi(\lambda)=|\lambda I-A|= (\lambda-\lambda_1)^{\tau_1}\dots(\lambda-\lambda_i)^{\tau_i}\dots(\lambda-\lambda_\rho)^{\tau_\rho}$. For $\lambda=\lambda_i$ compute:

$r_i^k = \text{rank}(A-\lambda_i I)^k$, $k=0,1,2,\ldots,q_i$, where q_i=index of annihilation of $A-\lambda_i I$. Note that $r_i^o = n$, $r_i^o - r_i^1 = d$=geomentric multiplicity of λ_i.

We define as <u>Weyr</u> characteristic the following set

$$\{r_i^o - r_i^1,\ r_i^1 - r_i^2, \ldots,\ r_i^{q_i-1} - r_i^{q_i}\} \triangleq \mathcal{W} \tag{2.77}$$

By using the Weyr characteristic we can construct the following Ferrer's diagram. In each row of this diagram we put as many asterisks as the numbers $r_i^o - r_i^1$, $r_i^1 - r_i^2, \ldots$ as indicated bellow:

$$\begin{array}{l}
r_i^o - r_i^1 \rightarrow * \; * \; \cdots \; * \; * \\
r_i^1 - r_i^2 \rightarrow * \; * \; \cdots \; * \\
\quad \vdots \\
r_i^{q_i-1} - r_i^{q_i} \rightarrow * \; * \; \cdots \\
\qquad\qquad\quad \uparrow \; \uparrow \qquad \uparrow \;\; \uparrow \\
\qquad\qquad\quad \theta_1 \, \theta_2 \qquad \theta_{d-1} \, \theta_d
\end{array} \tag{2.78}$$

The set $\{\theta_1, \theta_2, \ldots, \theta_{d-1}, \theta_d\} \triangleq \mathcal{J}$ which is the set of numbers of asteriscs in the columns of the Ferrer's diagram is the <u>Segrè</u> characteristic. The <u>Segrè</u> characteristic $\{\theta_1, \theta_2, \ldots, \theta_d\}$ defines the dimensions of each of the d independent chains.

2.4.4. Jordan canonical form

Let $T \in L(\mathbb{C}^n)$. Then [Hir. & Smale -1] there are unique operators S, N on $\mathbb{C}^n$ such that $T=S+N$, where $SN=NS$ and S is diagonalizable (semisimple) and N is nilpotent. In the case where $\sigma(T)=\{\lambda\}$ we have $S=\lambda I$ and hence $T=\lambda I+N$.

We start with an operator $T \in L(\mathbb{C}^n)$ that has only one eigenvalue λ, ($\sigma(T)=\{\lambda\}$); in that case $T=\lambda I+N$ with N nilpotent. Also there is a basis $\mathcal{B}=\{\underline{\beta}_1, \underline{\beta}_2, \ldots, \underline{\beta}_n\}$ that gives N a matrix representation in nilpotent canonical form A. That means that A is composed of diagonal blocks, each of which is an elementary nilpotent matrix H_{k_i}, $i=\underset{\sim}{\rho}$. So, we have:

$$[T]_{\mathcal{B}} = \lambda I_n + A = \text{block diag.}\{\lambda I_{k_1}+H_{k_1},\dots,\lambda I_{k\rho}+H_{k\rho}\} \tag{2.79}$$

The blocks making up λI_n+A are called elementary Jordan matrices, or elementary. λ-block. A matrix of the form (2.79) is called a Jordan matrix belonging to λ, or briefly, a Jordan λ-block.

Consider next an operator $T:\mathbb{C}^n\to\mathbb{C}^n$ whose distinct eigenvalues are $\lambda_1,\lambda_2,\dots,\lambda_m$; Then $\mathbb{C}^n=\oplus\sum_{i=1} U_i$, where $U_i=N_g(T,\lambda_i)$ is the generalised λ_k-eigenspace, $k=1,2,\dots,m$. We know that $T/U_k=\lambda_k I+N_k$, $k=1,2,\dots,m$ with N_k nilpotent. We give U_k a basis $\mathcal{B}_k$ which gives T/U_k a Jordan matrix belonging to λ_k. The basis $\mathcal{B}=\bigcup_{i=1}^{k}\mathcal{B}_i$ of $\mathbb{C}^n$ gives T a matrix representation of the form:

$$[T]_{\mathcal{B}} = C = \text{diag.}\{C_1,C_2,\dots,C_m\} \tag{2.80}$$

where each C_k is a Jordan matrix belonging to λ_k. Thus C is composed of diagonal blocks, each of which is an elementary Jordan matrix C_i. The matrix C is called the Jordan form of T.

It is easy to prove that similar operators have the same Jordan forms (perhaps with rearranged λ-blocks). For if $PT_oP^{-1}=T_1$, then P maps each generalised λ-eigenspace of T_o isomorphically onto the generalised λ-eigenspace of T_1; hence the Jordan λ-blocks are the same for T_o and T_1.

Note that the Jordan canonical form may be constructed from the set of Segrè characteristics which correspond to all distrinct eigenvalues. In fact, every number, k, in the Segrè characteristic $S(\lambda)$ of the eigenvalue λ defines a kxk elementary Jordan matrix that corresponds to the eigenvalue λ.

2.5. Matrix pencils

Let $(F,G)\in\mathbb{R}^{mxn}\times\mathbb{R}^{mxn}$ and $(s,\hat{s})$ be a pair of indeterminates. The polynomial matrix $sF-\hat{s}G\in\mathbb{R}^{mxn}[s,\hat{s}]$ is defined as the homogeneous matrix pencil of the pair (F,G). Clearly, $sF-\hat{s}G$ is a matrix over the ring $\mathbb{R}[s,\hat{s}]$ (polynomials in $(s,\hat{s})$ with coefficients from $\mathbb{R}$), but it may be also viewed as a matrix over the rings $\mathbb{R}(s)[\hat{s}]$, or $\mathbb{R}(\hat{s})[s]$. On the set $\mathcal{L}_{m,n}(s,\hat{s})$ we may define the following notions of equivalence.

Definition (2.23): Let $sF-\hat{s}G$, $sF'-\hat{s}G' \in \mathcal{L}_{m,n}(s,\hat{s})$ and let $R(s,\hat{s}) \in \mathbb{R}^{m\times m}(s,\hat{s})$, $Q(s,\hat{s}) \in \mathbb{R}^{n\times n}(s,\hat{s})$, $|R(s,\hat{s})|=c_1(s,\hat{s})\neq 0$, $|Q(s,\hat{s})|=c_2(s,\hat{s})\neq 0$ for which

$$R(s,\hat{s})\{sF-\hat{s}G\}Q(s,\hat{s}) = sF'-\hat{s}G' \qquad (2.81)$$

(i) If $R(s,\hat{s})$, $Q(s,\hat{s})$ are defined over $\mathbb{R}(\hat{s})[s]$ ($\mathbb{R}(s)[\hat{s}]$) and $c_1(s,\hat{s})$, $c_2(s,\hat{s}) \in \mathbb{R}(\hat{s})-\{0\}$ ($\mathbb{R}(s)-\{0\}$), then $sF-\hat{s}G$, $sF'-\hat{s}G'$ are said to be $\mathbb{R}(\hat{s})[s]$-equivalent ($\mathbb{R}(s)[\hat{s}]$-equivalent) and shall be denoted by $(sF-\hat{s}G)\mathcal{E}_{\mathbb{R}(\hat{s})[s]}(sF'-\hat{s}G')$ $((sF-\hat{s}G)\mathcal{E}_{\mathbb{R}(s)[\hat{s}]}(sF'-\hat{s}G'))$.

(ii) If $R(s,\hat{s})$, $Q(s,\hat{s})$ are defined over $\mathbb{R}[s,\hat{s}]$ and $c_1(s,\hat{s})$, $c_2(s,\hat{s}) \in \mathbb{R}-\{0\}$ then $sF-\hat{s}G$, $sF'-\hat{s}G'$ are said to be $\mathbb{R}[s,\hat{s}]$-equivalent and shall be denoted by $(sF-\hat{s}G)\mathcal{E}_{\mathbb{R}[s,\hat{s}]}(sF'-\hat{s}G')$.

(iii) If $R(s,\hat{s})$, $Q(s,\hat{s})$ are defined over $\mathbb{R}$ and c_1, $c_2 \in \mathbb{R}-\{0\}$, then $sF-\hat{s}G$, $sF'-\hat{s}G'$ are called strict equivalent [Gan. -1] and shall be denoted by $(sF-\hat{s}G)\mathcal{E}_s(sF'-\hat{s}G')$.

□

It is readily shown that the above relations are equivalence relations. The symbols $\mathcal{E}_{\mathbb{R}(\hat{s})[s]}$, $\mathcal{E}_{\mathbb{R}(s)[\hat{s}]}$, $\mathcal{E}_{\mathbb{R}[s,\hat{s}]}$ and $\mathcal{E}_s$ will be used for these equivalence relations.

Definition (2.24): The Smith form over $\mathbb{R}[s,\hat{s}]$ of the pencil $sF-\hat{s}G \in \mathcal{L}_{m,n}(s,\hat{s})$ is defined as the matrix

$$S(s,\hat{s}) = \begin{array}{cc} \left[\begin{array}{c|c} S^*(s,\hat{s}) & 0 \\ \hline 0 & 0 \end{array}\right] & \begin{array}{l} \updownarrow \rho \\ \updownarrow m-\rho \end{array} \\ \begin{array}{cc} \xleftrightarrow{\quad} & \xleftrightarrow{\quad} \\ \rho & n-\rho \end{array} & \end{array} \qquad (2.82a)$$

where $\rho=\mathrm{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}$ and $S^*(s,\hat{s})=\mathrm{diag.}\{f_i(s,\hat{s}),\ i\in \underset{\sim}{\rho}\}\in\mathbb{R}^{\rho\times\rho}[s,\hat{s}]$; the $f_i(s,\hat{s})$ are the <u>invariant polynomials over $\mathbb{R}[s,\hat{s}]$</u> of $sF-\hat{s}G$. If $\{d_i(s,\hat{s}),\ i=0, 1,\ldots,\rho,\ d_o(s,\hat{s})=1\}$ is the set of deterninantal divisors of $sF-\hat{s}G$ ($d_i(s,\hat{s})$ is the g.c.d. of all minors of order i), then the invariant polynomials $f_i(s,\hat{s})$ may be defined by the standard Smith algorithm [Tur. & Ait. -1] as follows:

$$f_i(s,\hat{s}) = d_i(s,\hat{s})/d_{i-1}(s,\hat{s}),\ i=1,\ldots,\rho,\quad d_o(s,\hat{s})=1 \qquad (2.82b)$$

The polynomials $f_i(s,\hat{s})$ are monic over $\mathbb{R}[s,\hat{s}]$ and it can be shown that $f_i(s,\hat{s})$ divides $f_{i+1}(s,\hat{s})$ $(f_i(s,\hat{s})/f_{i+1}(s,\hat{s}))$ for $\forall\ i\leq\rho-1$.

☐

The <u>Smith form $S^{\hat{s}}(s,\hat{s})$ over $\mathbb{R}(s)[\hat{s}]$</u> (<u>Smith form $S^{s}(s,\hat{s})$</u> over $\mathbb{R}(s)[\hat{s}]$) of $sF-\hat{s}G$ has the same form as $S(s,\hat{s})$ in (2.82a) except that the invariant polynomials are defined in terms of the determinantal divisors which are monic over $\mathbb{R}(\hat{s})[s]$ ($\mathbb{R}(s)[\hat{s}]$). Some interesting observations on the relationships between the various Smith forms have led to the introduction of the motion of dual pencils [Kar. & Hay -1,2,3]. This topic will be discussed in following chapter. In the following we give a summary of the well known results of matrix pencil theory under strict equivalence [Gan. -1][Tur. & Ait. -1].

<u>Definition (2.23)</u>: The pencil $sF-\hat{s}G$ is said to be <u>regular</u> if $F, G\in\mathbb{R}^{n\times n}$ and $\rho=\mathrm{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}=n$; in all other cases, i.e. $m=n$ and $\rho<n$, $m<n$ or $m>n$, it will be called <u>singular</u>.

☐

Note that the terms given for the pencils, will be also used for the pairs (F,G) which "generate" the pencil $sF-\hat{s}G$. In the following the single variable pencil $sF-G$ is used; the summary of the results is based on the treatment given in [Gan. -1].

By factorising the invariant polynomials $f_i(s,\hat{s})$ of $S(s,\hat{s})$ into powers of homogeneous polynomials irreducible over $\mathbb{C}$, we obtain the set of <u>elementary divisors</u> (e.d.) of the pencil $sF-\hat{s}G$; these are of the following type: $s^p, \hat{s}^q$, and pairs of complex conjugate e.d. $(s-\alpha\hat{s})^{\tau}$, $(s-\bar{\alpha}\hat{s})^{\tau}$, $\alpha,\bar{\alpha} \in \mathbb{C}$. For the pencil sF-G we define: e.d. of the type $\hat{s}^q$ are called <u>infinite elementary divisors</u> (i.e.d.), e.d. of the type s^p are called <u>zero elementary divisors</u> (z.e.d.), and e.d. of the type $(s-\alpha)^{\tau}$ are called <u>non zero finite elementary divisors</u> (nz.f.e.d.). Whenever, there is no distinction between the e.d. of the type s^p, $(s-\alpha)^{\tau}$, $\alpha\neq 0$, they will be referred to as as <u>finite elementary divisors</u> (f.e.d.). The set of all i.e.d., z.e.d., nz.f.e.d. and f.e.d. of sF-G will be denoted in short by {i.e.d.}, {z.e.d.}, {nz.f.e.d.}, {f.e.d.} respectively. whenever there is no ambiguity about the pencil sF-G on which these sets are defined. In the following $L^r_{n,n}$, $L^s_{m,n}$ shall denote the sets of regular, singular pencils of dimensions nxn, mxn respectively. An element sF-G, of the above sets, will be denoted in short by L. Note that strict equivalence have been defined over the real s; however it may also be defined over the complex numbers. The <u>real strict equivalence</u> will be denoted by $E^{\mathbb{R}}_S$ and the <u>complex strict equivalence</u> by $E^{\mathbb{C}}_S$; of course $E^{\mathbb{C}}_s$ is defined over sets of complex pencils. By $E^{\mathbb{R}}_S(L)$, $E^{\mathbb{C}}_S(L)$ we shall denote the corresponding equivalence classes.

The classical theory of matrix pencils deals with the study of invariant and canonical forms of the strict equivalence classes $E^{\mathbb{R}}_S(L)$, $E^{\mathbb{C}}_S(L)$, when $L \in L^r_{n,n}$ or $L \in L^s_{m,n}$. The results are presented below for single variable pencils sF-G; for homogeneous pencils sF-$\hat{s}$G, the results are similar (homogenise the results stated for sF-G).

<u>Theorem (2.7)</u>: The map $f:L^r_{n,n} \to$ {f.e.d.}x{i.e.d.} is a complete invariant for $E^{\mathbb{R}}_S(L)$, $E^{\mathbb{C}}_S(L)$, $L \in L^r_{n,n}$.

□

Note that for $E^{\mathbb{R}}_S$-equivalence the invariant polynomials are factorised over $\mathbb{R}$ and thus {f.e.d.} is made up from irreducible over $\mathbb{R}$ e.d. of the type $(s-\alpha)^{\tau}, \alpha \in \mathbb{R}$, $(s^2-\beta s-\gamma)^k$, $\alpha,\beta,\gamma \in \mathbb{R}$. The existence of a complete set of invariants for pencils $sF-G \in \mathcal{L}^{r}_{n,n}$ implies the existence of a canonical form, known as Weierstrass canonical form [Gan. -1][Tur. & Ait. -1].

□

Theorem (2.8): The equivalence class $\mathcal{E}^{\mathbb{C}}_{S}(L)$, $L \in \mathcal{L}^{r}_{n,n}$ is characterised by a canonical element, sF_w-G_w, the complex Weierstrass canonical form, defined by

$$sF_w-G_w = \text{diag.}\{\dots;sH_q-I_q;\dots;sI_\tau-J_\tau(\lambda);\dots\} \tag{2.83}$$

where $sI_\tau-J_\tau(\lambda)$ is the Jordan canonical block associated with the f.e.d. $(s-\lambda)^\tau$ and sH_q-I_q is a canonical block associated with an i.e.d. $\hat{s}^q$ (H_q is an elementary qxq nilpotent matrix, whose elements in the first superdiagonal are one, whereas the remaining elements are all zero).

□

A real Weierstrass form, $sF'_w-G'_w$, may be constructed by appropriate modification of the blocks in (2.83), as it will be shown below; $sF'_w-G'_w$ will be then a canonical form of the $\mathcal{E}^{\mathbb{R}}_{S}(L)$ class. Thus, let $\alpha=-\sigma-j\omega$, $\bar{\alpha}=-\sigma+j\omega \in \mathbb{C}$ and let $q(s)=(s-\alpha)^\tau(s-\bar{\alpha})^\tau=(s^2-\beta s-\gamma)^\tau$ be the quadratic real elementary divisor that corresponds to the pair of complex conjugate e.d. $(s-\alpha)^\tau$, $(s-\bar{\alpha})^\tau$. Using standard results from [Tur. & Ait. -1], we may construct two different real Weierstrass forms. Thus, to $q(s)$ we may associate the canonical blocks $C_\tau(\beta,\gamma)$, $D_\tau(\sigma,\omega)$ of dimensions $2\tau x 2\tau$ where

$$C_\tau(\beta,\gamma) \triangleq \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \gamma & \beta & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \gamma & \beta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \dots & \gamma & \beta \end{bmatrix} \tag{2.84a}$$

and

$$D_\tau(\sigma,\omega) = \begin{bmatrix} W & I_2 & 0 & \dots & 0 & 0 \\ 0 & W & I_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & W & I_2 \\ 0 & 0 & 0 & & 0 & \end{bmatrix}, \quad W = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} \tag{2.84b}$$

It can be proved that under $\mathcal{E}_S^{\mathbb{R}}$-equivalence, $sF-G \in L^r_{n,n}$ may be reduced to a real canonical form, of the type (2.83), where instead of a pair of blocks $sI_\tau - J_\tau(\alpha)$, $sI_\tau - J_\tau(\bar{\alpha})$, the blocks $C_\tau(\beta,\gamma)$, or $D_\tau(\sigma,\omega)$ are used; the first of those two canonical forms will be referred to as the Real Weierstrass canonical form, whereas the second will be called the Real spectral Weierstrass canonical form [Kar. & Hay. -2].

Unlike the case of regular pencils, however, the characterisation of the $\mathcal{E}_S^{\mathbb{R}}(L)$ class, $L \in L^s_{m,n}$, apart from the set of the determinantal divisors requires the definition of additional sets of invariants, the minimal indices. Thus assume that $m \neq n$ and that $\rho = \text{rank}_{\mathbb{R}(s)}\{sF-G\} < \min\{m,n\}$. Then the equations

$$(sF-G)\underline{x} = \underline{0}\ , \qquad y^t(sF-G) = \underline{0}^t \tag{2.85}$$

have solutions in $\underline{x}$ and $\underline{y}$, which are vectors in the rational vector spaces $N_r(s) = N_r\{sF-G\}$ and $N_\ell(s) = N_\ell\{sF-G\}$ respectively.

Let $p = \dim N_r(s)$, $t = \dim N_\ell(s)$. It is known [For. -1] that $N_r(s)$ and $N_\ell(s)$, as rational vector spaces, are spanned by minimal polynomial bases $\{\underline{x}_i(s),\ i \in \underset{\sim}{p}\}$ and $\{\underline{y}_j^t(s),\ j \in \underset{\sim}{t}\}$ correspondingly, of minimal degrees $\{\varepsilon_1 = \dots = \varepsilon_g = 0 < \varepsilon_{g+1} \leq \dots \leq \varepsilon_p\}$ and $\{\zeta_1 = \dots = \zeta_{g'} = 0 < \zeta_{g'+1} \leq \dots \leq \zeta_t\}$ respectively. The set of minimal indices $\{\varepsilon_i\}$ and $\{\zeta_j\}$ are known [Gan. -1] as column minimal indices (c.m.i.) and row minimal indices (r.m.i.) of sF-G respectively. Note that if the homogeneous pencil, $sF-\hat{s}G$, is considered, then the same set of minimal indices is defined for the two rational vector spaces N_r, N_ℓ, which are now defined over $\mathbb{R}(s,\hat{s})$ [Tur. & Ait. -]. The sets of c.m.i., r.m.i. of sF-G will be denoted in short by {c.m.i.}, {r.m.i.} correspondingly.

Theorem (2.9): The map $f:L^{s}_{m,n} \to$ {f.e.d.}x{i.e.d.}x{c.m.i.}x{r.m.i.} is a complete invariant for $E^{\mathbb{R}}_{S}(L)$, $E^{\mathbb{C}}_{S}(L)$, $L \in L^{s}_{m,n}$.

□

The existence of a complete set of invariants for pencils $sF-G \in L^{s}_{m,n}$ implies the existence of a canonical form, known as Kronecker canonical form [Gant. -1].

Theorem (2.10): The equivalence class $E^{\mathbb{C}}_{S}(L)$, $L \in L^{s}_{m,n}$ is characterised by a canonical element, sF_k-G_k, the complex Kronecker canonical form, defined by

$$sF_k-G_k = \text{quasi diag.}\{0_{h,g};L_{\varepsilon_{g+1}},\dots,L_{\varepsilon_p};L^{t}_{n_{h+1}},\dots,L^{t}_{n_t};sF_w-G_w\} \qquad (2.86)$$

where sF_w-G_w is the complex Weierstrass canonical form, associated with the f.e.d. and i.e.d. of the pencil, $0_{h,g}$ is a zero block parametrized by the g zero c.m.i. and h zero r.m.i. and L_{ε_i} and $L^{t}_{n_j}$ are blocks corresponding to non zero c.m.i. and r.m.i. respectively, of the type

$$L_\xi = \underbrace{\begin{bmatrix} s & -1 & 0 & \dots & 0 & 0 \\ 0 & s & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s & -1 \end{bmatrix}}_{\longleftarrow\ \xi+1\ \longrightarrow} \updownarrow \xi \qquad \xi=\varepsilon_i \text{ or } n_j \quad . \qquad (2.87)$$

The real Kronecker canonical form, $sF'_k-G'_k$, may be defined by

$$sF'_k-G'_k = \text{quasi diag.}\{0_{h,g};L_{\varepsilon_{g+1}},\dots,L_{\varepsilon_p};L^{t}_{n_{h+1}},\dots,L^{t}_{n_t};sF'_w-G'_w\} \qquad (2.88)$$

where $sF'_w-G'_w$ is the real Weierstrass form.

2.6. Topological bankground

The norm for a vector $\underline{x} \in \mathbb{R}^n$ will be denoted by $\|\underline{x}\|$ and satisfies the following relations:

(i) $\|\underline{x}\| \geq 0 \quad \forall\, \underline{x} \in \mathbb{R}^n$ unless $\underline{x}=\underline{0}$

(ii) $\|k\underline{x}\| = |k| \cdot \|\underline{x}\| \quad \forall\, k \in \mathbb{C},\ \underline{x} \in \mathbb{R}^n$ (2.89)

(iii) $\|\underline{x}+\underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\| \quad \forall\, \underline{x},\underline{y} \in \mathbb{R}^n$

Also from (iii) we have $\|\underline{x}-\underline{y}\| \geq \big| \|\underline{x}\| - \|\underline{y}\| \big| \quad \forall\, \underline{x},\underline{y} \in \mathbb{R}^n$.

We use three vector norms. They are defined by:

$$\|\underline{x}\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{1/p} \qquad (p=1,2,\infty) \tag{2.90}$$

where $\underline{x}=[x_1,x_2,\ldots,x_n]^t \in \mathbb{R}^n$ and $\|\underline{x}\|_\infty$ is interpreted as $\max|x_i|$, $i=1,2,\ldots,n$. The norm $\|\underline{x}\|_2$ is the Euclidean length of the vector $\underline{x}$. Corresponding to any vector norm a non-negative quantity, defined by $\operatorname{Sup}_{\underline{x}\neq\underline{0}} \|A\underline{x}\| / \|\underline{x}\|$, may be associated with any matrix A.

From (ii) of (2.89) we see that this is equivalent to $\sup_{\|\underline{x}\|=1} \|A\underline{x}\|$.

Also for $\|A\| \triangleq \sup_{\|x\neq 0\|} \|A\underline{x}\| / \|\underline{x}\| = \sup_{\|\underline{x}\|=1} \|A\underline{x}\|$ we have:

$$\|A\underline{x}\| \leq \|A\| \cdot \|\underline{x}\| \tag{2.91}$$

Matrix and vector norms for which (2.91) is true for all A and $\underline{x}$ are said to be compatible.

The matrix norm subordinate to $\|\underline{x}\|_p$ is denoted by $\|A\|_p$.

These norms satrisfy the relations

$$\|A\|_1 = \max_j \sum_i |\alpha_{ij}|$$

$$\|A\|_\infty = \max_i \sum_j |\alpha_{ij}| \tag{2.92}$$

$$\|A\|_2 = [\lambda_{\max}(A^H A)]^{1/2}$$

$\|\cdot\|_2$ is called the spectral norm or the bound norm.

Another norm which is compatible with the Euclidean vector norm is the <u>Frobenious</u> or <u>Hilbert-Schmidt</u> norm:

$$\|A\|_F \triangleq (\text{trace}(A^H A))^{1/2} = [\sum_{i,j} |\alpha_{ij}|^2]^{1/2} \tag{2.93}$$

For any mxn matrix A, with m≥n, we have:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{\text{rankA}} \cdot \|A\|_2 \leq \sqrt{n} \cdot \|A\|_2 \tag{2.94}$$

when $A=A^H$ then $\|A\|_2 = \max\{|\lambda_1|,|\lambda_2|,\ldots,|\lambda_n|\}$=spectral radious of A and $\|A\|_F = [\sum_{i=1}^{n} \lambda_i^2]^{1/2}$.

The next proposition says that a sequence of unitary transformations cannot change either norm.

□

<u>Proposition (2.10)</u>: For any A mxn matrix we have $\|PAQ\|_2 = \|A\|_2$, $\|PAQ\|_F = \|A\|_F$ if and only if P, Q are orhonormal, i.e. $P^H P = PP^H = I_m$, $Q^H Q = QQ^H = I_n$. □

<u>Definitions 2.24 [Kato -1]</u>: Let $X, Y \subset \mathbb{C}^n$ be subspaces. The gap between X and Y is the number:

$$\gamma(X,Y) \triangleq \max\left[\sup_{\substack{\|\underline{x}\|=1,\\ \underline{x}\in X}} \inf_{\underline{y}\in Y} \|\underline{x}-\underline{y}\|, \quad \sup_{\substack{\|\underline{y}\|=1,\\ \underline{y}\in Y}} \inf_{\underline{x}\in X} \|\underline{y}-\underline{x}\|\right] \tag{2.95}$$

The gap function is not quite a metric, for it does not satisfy the triangle inequality, however the neighborhoods $N(X,\varepsilon)$, $\varepsilon>0$ defined by:
$N(X,\varepsilon) \triangleq \{Y: \gamma(X,Y)<\varepsilon\}$ form a basis for a topology on the set of all subspaces of $\mathbb{C}^n$.

In the case where the norm used in the definition of gap is the 2-norm the gap function is a metric.

More important for our purposes is the following theorem that relates the gap between X and Y to the projectors onto X and Y.

Theorem (2.11)[Kato -1]: Let $P_{\mathcal{X}}$, $P_{\mathcal{Y}}$ are the orthogonal projectors onto the subspaces $\mathcal{X}$ and $\mathcal{Y}$ of $\mathbb{C}^n$. If the 2-norm is used to define the gap in previous definition then:

$$\gamma(\mathcal{X},\mathcal{Y}) = ||P_{\mathcal{X}}-P_{\mathcal{Y}}||_2$$

□

CHAPTER 3:

Generalised Autonomous Differential Systems, Matrix Pencils and Linear Systems

CHAPTER 3: GENERALISED AUTONOMOUS DIFFERENTIAL SYSTEMS MATRIX PENCILS AND LINEAR SYSTEMS

3.1 Introduction

The aim of this chapter is to "set the scene" for the theory that we will develop in this thesis. It will be shown that problems of regular and extended state space theory may be described in quite a natural way in terms of generalised autonomous differential systems. It is because of this unified description, that matrix pencil theory plays an instrumental role in the study of algebraic, geometric and dynamic properties of linear systems. The generalised autonomous differential systems provide a natural setting for defining different notions of duality; the notion of "integrator-differentiator" type of duality, motivates the study of a new type of equivalence on matrix pencils, namely the "Bilinear-Strict equivalence". The link between the algebraic structure of matrix pencils with the structure of the subspaces of the domain and codomain of a pair (F,G), motivates the need for the developement of a geometric theory of matrix pencils; detaching this theory from its algebraic basis is of considerable importance, since extensions of the theory to pencils of more general operators is easier in a geometric, rather than an algebraic context. The equivalence between the dynamic, geometric characterisations of invariant subspaces of the geometric theory [Won. -1], [Will. -1] and the algebraic characterisation provided by the restriction pencil [Kar. -1], [Jaf. & Kar. -1], indicates that the subspaces of the domain of (F,G) may be characterised dynamically.

The chapter is structured as follows: In Section (3.2) we will briefly recall some of the basic concepts from state space theory, in particular those related to the dynamic, geometric characterisation of subspaces of the state space. In Section (3.3) it is shown how the generalised autonomous differential systems, and thus matrix pencils, arise in the theory of regular and extended state space linear systems. Section (3.4) reviews the fundamental duality notions on autonomous differential descriptions. Finally, Section (3.5) explains the motivation behind the theory developed in the subsequent chapters.

3.2 Background concepts from linear geometric theory

Consider the dynamical system whose evolution in time is modelled by the linear time invariant differential and algebraic equations

$$S(A,B,C,D): \quad \begin{aligned} \dot{\underline{x}}(t) &= A\underline{x}(t)+B\underline{u}(t) \quad t\geq 0 && (3.1a) \\ \underline{y}(t) &= C\underline{x}(t)+D\underline{u}(t) && (3.1b) \end{aligned}$$

where $\underline{u}(\cdot): \mathbb{R}_+\to U$, $\underline{x}(\cdot): \mathbb{R}_+\to X$, $\underline{y}(\cdot): \mathbb{R}_+\to Y$ and U,X,Y are real linear rector spaces, with dim $U=l$, dim $X=n$, dim $Y=m$. Here X is the state space, Y the output space and U the input space. A,B,C,D are linear mappings defined by A: $X\to X$, B: $U\to X$, C: $X\to Y$, D: $U\to Y$ and they are referred to as state-, input-state, state-output, input-output maps respectively. $\underline{u}(\cdot)$ may be piecewise continuous, C^{∞}, or distributional; for the study of most properties the piecewise continuous assumption for $\underline{u}(\cdot)$ is sufficient. The linear spaces X,U,Y are isomorphic to $\mathbb{R}^n$, $\mathbb{R}^l$, $\mathbb{R}^m$ and thus A,B,C,D may be represented in terms matrices of $\mathbb{R}^{nxn}$, $\mathbb{R}^{nxl}$, $\mathbb{R}^{mxn}$, $\mathbb{R}^{mxl}$ correspondingly. Throughout the thesis we shall consider the matrix representations of A,B,C,D maps and we shall use the same notation. It will be also assumed that rank(B)=l and rank(C)=m.

The S(A,B,C,D) model is usually referred as a regular state space model, to distinguish from a more general model defined subsequently. Whenever the symbols S(A), S(A,B), S(A,B,C) are used, that means we consider the systems described by the corresponding maps. A more general than the S(A,B,C,D) model will be also considered, i.e.

$$S_e(E,A',B',C',D'): \quad \begin{aligned} E\,\dot{\underline{x}}(t) &= A'\underline{x}(t)+B'\underline{u}(t) \quad t\geq 0_- && (3.2a) \\ \underline{y}(t) &= C'\underline{x}(t)+D'\underline{u}(t) && (3.2b) \end{aligned}$$

where $\underline{u}(\cdot)\colon \mathbb{R}_+\to\mathbb{R}^1$, $\underline{x}(\cdot)\colon \mathbb{R}_+\to\mathbb{R}^n$, $\underline{y}(\cdot)\colon \mathbb{R}_+\to\mathbb{R}^m$ and $E,A'\in\mathbb{R}^{nxn}$, $B'\in\mathbb{R}^{nx1}$, $C'\in\mathbb{R}^{mxn}$, $D'\in\mathbb{R}^{mx1}$. Note that E is singular (otherwise, this description may be reduced to the previous one) and it is known as extended state-space model [Ros. -2], [Verg. -1], or descriptor type model [Lue. -1]. There are essential differences between the system models $\mathcal{S}$ and $\mathcal{S}_e$, the most important one has to do with the nature of the solutions of $\mathcal{S}(A)$ and $\mathcal{S}_e(A)$ (see [Camp. -1,2]).

In the following we shall consider the $\mathcal{S}(A,B)$ system and shall give a brief summary of the important concepts of subspaces developed in the geometric thoery; for a full exposition see [Won. -1], [Wil. -1].

The system $\mathcal{S}(A)$: $\dot{\underline{x}}(t)=A\underline{x}(t)$ represents a linear flow [Hir. & Sm. -1]. An important notion in the study of flows is the notion of invariance. A subspace $\mathcal{V}\subset\mathcal{X}$ will be said to be dynamically invariant with respect to the flow $\mathcal{S}(A)$, if every initial condition $\underline{x}(o)=\underline{x}_o\in\mathcal{V}$ gives rise to a trajectory $\underline{x}(t)$ that lies entirely in $\mathcal{V}$. We will call $\mathcal{V}$ geometrically invariant, or A-invariant if $A\mathcal{V}\subset\mathcal{V}$. It is clear that the families of dynamically and geometrically invariant subspaces coincide. Suppose now that it is possible to change the dynamic behaviour of $\mathcal{S}(A)$ by some external mechanism. In particular assume that our system is given by

$$\mathcal{S}(A,B)\colon\ \ \dot{\underline{x}}(t) = A\underline{x}(t)+B\underline{u}(t) \tag{3.3}$$

Under the further assumption that we are allowed to use for $\underline{u}(t)$ linear functions of the current state $\underline{x}(t)$ (accessibility of the states for measurement) we may take $\underline{u}(t)=F\underline{x}(t)$ in (3.3), where F is a linear map from $\mathcal{X}$ into $\mathcal{U}$, called a state feedback. With this expression for $\underline{u}(t)$, we may change the dynamics of (3.3) to

$$\mathcal{S}'(A+BF)\colon\ \ \dot{\underline{x}}(t) = (A+BF)\underline{x}(t) \tag{3.4}$$

It is clear that the notions of dynamically invariant and geometrically invariant subspaces may be extended to the $\mathcal{S}(A,B)$ systems. These observations have led to the developement of the geometric theory of linear systems

[Bas. & Mar. -1], [Won. & Mor. -1], [Won. -1], [Will. -1]; the fundamental concepts and results are summarised below.

Consider the system $S(A,B)$, $\underline{x}(o)=\underline{x}_o$ denote an initial condition and let $\underline{u}(t)$ be a control input. The resulting state trajectory is given by

$$\underline{x}(t) = e^{At}\underline{x}_o+\int_o^t e^{(t-\tau)}B\underline{u}(\tau)d\tau \tag{3.5}$$

The linear space of all absolutely continuous trajectories will be denoted by $\Sigma(A,B)$ and it is formally defined by

$$\Sigma(A,B) \overset{\Delta}{=} \{\underline{x}: \mathbb{R}\to X: x \text{ is absolutely continuous} \\ \text{and } \underline{\dot{x}}(t)-A\underline{x}(t) \in \text{im } B \text{ a.c.}\} \tag{3.6}$$

In the following, if $F: X\to U$ is a mapping we shall often denote $A_F \overset{\Delta}{=} A+BF$. The space $\text{Im}B\subset X$ shall be denoted by B.

<u>Definition (3.1)</u>: (i) A subspace V of X will be called <u>controlled invariant</u> subspace, if $\forall\underline{x}_o\in V$, $\exists\underline{x}\in\Sigma(A,B)$ such that $\underline{x}(o)=\underline{x}_o$ and $\underline{x}(t)\in V$, $\forall t\in\mathbb{R}_+$.
(ii) A subspace R of X will be called a <u>controllabity subspcace</u> (c.s.) if $\forall\underline{x}_o$, $\underline{x}_1\in R$, $\exists\tau>0$ and $\underline{x}\in\Sigma(A,B)$ such that $\underline{x}(o)=\underline{x}_o$, $\underline{x}(\tau)=\underline{x}_1$ and $\underline{x}(t)\in R$, $\forall t\in\mathbb{R}_+$.

□

Controlled invariant subspaces are also called (A,B)-invariant subspaces, or A(mod B)-invariant subspaces. The classes of all controlled invariant, controllability subspaces defined on $S(A,B)$ shall be denoted by $V(A,B)$, $R(A,B)$ respectively. The notion of c.s. is an extension of the standard notion of controllability [Kal. -1]; the meaning of a c.s. is that it is possible to travel between any two points of the subspace, moving along a trajectory that lies entirely in that subspace. It is clear that c.s. are always controlled invariant subspaces, and thus $R(A,B)\subset V(A,B)$. In the following, if B_1 is a subspace of B and $F: X\to U$ is a mapping, we will demote $\langle A_F/B_1\rangle \overset{\Delta}{=} B_1+A_FB_1+\ldots+A_F^{n-1}B_1$. The classes $V(A,B)$, $R(A,B)$ have been characterised geometrically [Bus. & Mar. -1], [Won. & Mor. -1] by the following results:

Proposition (3.1): The following statements are equivaltent:

(i) $\mathcal{V}\in\{\mathcal{V}(A,B)\}$,

(ii) $\exists F: \mathcal{X}\to\mathcal{U}$ such that $A_F\mathcal{V}\subset\mathcal{V}$

(iii) $A\mathcal{V}\subset\mathcal{V}+\mathcal{B}$

□

Proposition (3.2): The following statements are equivalent:

(i) $\mathcal{R}\in\{\mathcal{R}(A,B)\}$

(ii) $\exists$a subspace $\mathcal{B}'\subset\mathcal{B}$ and a mapping $F: \mathcal{X}\to\mathcal{U}$ such that $\mathcal{R}=\langle A_F/\mathcal{B}'\rangle$

(iii) $\exists$a mapping $F: \mathcal{X}\to\mathcal{U}$ such that $\mathcal{R}=\langle A_F/\mathcal{B}\cap\mathcal{R}\rangle$

□

For a proof of these results see [Won. -1]. It may be readily proved that the classes $\mathcal{V}(A,B)$ and $\mathcal{R}(A,B)$ are closed under the operation of subspace addition. Because of this latter property for every subspace $\mathcal{K}\subset\mathcal{X}$, there is a supremal (maximal) controlled invariant subspace and a supremal (maximal) c.s. which are contained in $\mathcal{K}$; these subspaces will be denoted by $\mathcal{V}^*(\mathcal{K})$ and $\mathcal{R}^*(\mathcal{K})$ respectively. The subspaces $\mathcal{V}^*(\mathcal{K})$ and $\mathcal{R}^*(\mathcal{K})$ may be computed by the following algorithms (see [Won. -1]).

Controlled invariant subspace algorithm: Let $\mathcal{K}\subset\mathcal{X}$ be a subspace. Consider the following sequence of subspaces

$$\mathcal{V}_{\mathcal{K}}^{o} = \mathcal{X}, \quad \mathcal{V}_{\mathcal{K}}^{\mu+1} = \mathcal{K}\cap A^{-1}(\mathcal{V}_{\mathcal{K}}^{\mu}+\mathcal{B}), \quad \mu\geq 0 \tag{3.7}$$

This sequence is non increasing and converges to $\mathcal{V}^*(\mathcal{K})$, the supremal (A,B)-(or controlled) invariant subspace contained in $\mathcal{K}$.

□

Controllability subspace algorithm: Let $\mathcal{K}\subset\mathcal{X}$ be a subspace. Consider the following sequence of subspaces

$$\mathcal{R}_{\mathcal{K}}^{o} = 0, \quad \mathcal{R}_{\mathcal{K}}^{\mu+1} = \mathcal{V}^*(\mathcal{K})\cap(A\mathcal{R}_{\mathcal{K}}^{\mu}+\mathcal{B}), \quad \mu\geq 0 \tag{3.8}$$

where $\mathcal{V}^*(\mathcal{K})$ is the supremal (A,B)-invariant subspace in $\mathcal{K}$. The above sequence is non decreasing and converges to $\mathcal{R}^*(\mathcal{K})$ the supremal controllability subspace contained in $\mathcal{K}$. □

The dynamic notions of controlled invariant and controllability subspaces have been extended by Willems [Will. -1] to the notions of almost controlled invariant subspace (a.c.i.s.) and almost controllability subspace (a.c.s.). The essence of the a.c.i.s. is that beginning a motion in it, one can stay arbitrarily close to it by choosing the input appropriately. In the same way, an a.c.s. has been defined as a subspace with the property that, starting in it, one can steer to an arbitrary point in the same subspace while staying arbitrarily close to that subspace. To formally define these two notions we need a measure of the distance of a point from a subspace. Thus, assume that $\mathcal{X}$ is a normed vector space with $\|\cdot\|$ norm. If $\mathcal{W}$ is a subspace of $\mathcal{X}$ and $\underline{x} \in \mathcal{X}$ we will define the distance of $\underline{x}$ to $\mathcal{W}$ to be

$$d(\underline{x},\mathcal{W}) \triangleq \inf_{\underline{x}' \in \mathcal{W}} \| \underline{x}-\underline{x}' \| \tag{3.9}$$

We may give the following definition:

Definition (3.2): (i) A subspace $\mathcal{V}_\alpha \subset \mathcal{X}$ is said to be an almost controlled invariant subspace (a.c.i.s.), if $\forall \underline{x}_o \in \mathcal{V}_\alpha$ and $\varepsilon>0$, $\exists x(t) \in \Sigma(A,B)$ such that $\underline{x}(o)=\underline{x}_o$ and $d(\underline{x}(t),\mathcal{V}_\alpha) \leq \varepsilon$, $\forall t \in \mathbb{R}_+$.

(ii) A subspace $\mathcal{R}_\alpha \subset \mathcal{X}$ is said to be an almost controllability subspace (a.c.s.), if $\forall \underline{x}_o, \underline{x}_1 \in \mathcal{R}_\alpha$, $\exists \tau>0$ such that $\forall \varepsilon>0$ $\exists \underline{x}(t) \in \Sigma(A,B)$ with the properties that $\underline{x}(o)=\underline{x}_o$, $\underline{x}(\tau)=\underline{x}_1$ and $d(\underline{x}(t),\mathcal{R}_\alpha) \leq \varepsilon$, $\forall t \in \mathbb{R}_+$. □

The families of a.c.i.s. and a.c.s. will be denoted by $\mathcal{V}_\alpha(A,B)$, $\mathcal{R}_\alpha(A,B)$ respectively. It is a trivial matter to verify that $\mathcal{R}(A,B) \subset \mathcal{V}(A,B) \subset \mathcal{V}_\alpha(A,B)$ and $\mathcal{R}(A,B) \subset \mathcal{R}_\alpha(A,B) \subset \mathcal{V}_\alpha(A,B)$. The families $\mathcal{V}_\alpha(A,B)$ and $\mathcal{R}_\alpha(A,B)$ are closed under the operation of subspace addition and thus it may immediately concluded that for every subspace $\mathcal{K}$ of $\mathcal{X}$, there is a supremal a.c.i.s. $\mathcal{V}_\alpha^*(\mathcal{K})$ and a supremal a.c.s. $\mathcal{R}_\alpha^*(\mathcal{K})$ contained in $\mathcal{K}$. A geometric characterisation of the families $\mathcal{V}_\alpha(A,B)$ and $\mathcal{R}_\alpha(A,B)$ has been given by Willems [Will. -1] and it is stated next.

<u>Proposition (3.3)</u>: (i) $\mathcal{R}_\alpha \in \mathcal{R}_\alpha(A,B)$, if and only if there is a mapping F: $\mathcal{X} \to \mathcal{U}$ and a chain $\mathcal{B} \supset \mathcal{B}_1 \supset \mathcal{B}_2 \supset \dots \supset \mathcal{B}_k$ such that

$$\mathcal{R}_\alpha = \mathcal{B}_1 + A_F\mathcal{B}_2 + \dots + A_F^{k-1}\mathcal{B}_k \qquad (3.10)$$

(ii) $\mathcal{V}_\alpha \in \mathcal{V}_\alpha(A,B)$, if and only if there exists $\mathcal{V} \in \mathcal{V}(A,B)$ and $\mathcal{R}_\alpha \in \mathcal{R}_\alpha(A,B)$ such that $\mathcal{V}_\alpha = \mathcal{V} + \mathcal{R}_\alpha$ □

The subspace $\mathcal{R}_\alpha^*(\mathcal{K})$ may be computed by a variation of the algorithm given for $\mathcal{R}^*(\mathcal{K})$ [Won. -1]. This algorithm is given below [Will. -1].

<u>Almost controllability subspace algorithm</u>: Let $\mathcal{K} \subset \mathcal{X}$ be a subspace. Consider the following sequence of subspaces

$$\mathcal{R}_{\alpha\mathcal{K}}^{o} = 0, \qquad \mathcal{R}_{\alpha\mathcal{K}}^{\mu+1} = \mathcal{K} \cap (A\mathcal{R}_{\alpha\mathcal{K}}^{\mu} + \mathcal{B}), \qquad \mu \geq 0 \qquad (3.11)$$

The sequence is non decreasing and converges to $\mathcal{R}_\alpha^*(\mathcal{K})$, the supremal almost controllability subspace contained in $\mathcal{K}$. □

The families $\mathcal{V}_\alpha(A,B)$, $\mathcal{R}_\alpha(A,B)$ have been also characterised dynamically in two different ways. The first, due to Willems [Will. -1], [Tren. -1] is making use of distributional inputs and the second due to Karcanius [Kar. -1], [MacF. & Kar. -1] is based on the problem of frequency transmission through subspaces [Kar. & Kouv. -2] of the state space of a linear system. The first of the above two approaches suggests that the families $\mathcal{V}_\alpha(A,B)$, $\mathcal{R}_\alpha(A,B)$ may be viewed as "exact" invariant subspaces when we allow the class of state trajectories to include not only absolutely continuous functions, but also distributions; this is achieved by introducing a convenient class of admissiable distributional inputs and thus allowing the definition of state trajectories satisfying certain initial conditions. It has been shown, that in order to give a characterisation of the $V_\alpha(A,B)$, $R_\alpha(A,B)$ families it is sufficient to consider an even smaller class of inputs, the class of inputs that are Bohl distributions [Will. -1]. This characterisation is similar to that given for controlled invariant subspaces and controllability subspaces (Definition (3.1)) with the only difference that $\Sigma(A,B)$ changes to $\Sigma_D(A,B)$, where the latter denotes the set of distributional solutions of $S(A,B)$. It is because of this characterisation and the equivalence of these new definitions to those given previously, that a $V_\alpha \in V_\alpha(A,B)$ is also referred as a <u>distributionally controlled invariant subspace</u> (d.c.i.s.) and an $R_\alpha \in R_\alpha(A,B)$ as a <u>distributionally controllability subspace</u> (d.c.s.). For a proper exposition of the topic see [Will. -1], [Tren. -1]. The second approach [Kar. -1] has been based on the "filter" properties of subspaces (type of frequencies that may be propagated through them) and was the first effort to extend the standard notions of the geometric theory. The characterisation of the subspaces of X was given in algebraic terms, i.e. by the type of invariants of a "restriction" pencil which may be associated with a subspace V. The matrix pencil approach has provided a simple characterisation of the subspaces of X and a simple method for the computation of $V^*(K)$, $V_\alpha{}^*(K)$, $R(K)$, $R_\alpha{}^*(K)$ in terms of the Kronecker canonical form. The equivalence of the subspaces introduced by algebraic means the subspaces defined by Willems [Will. -1] has been established by Jaffe and Karcanias [Jaf. & Kar. -1]. The matrix pencil approach forms the basis for the treatment given in this thesis and it will be briefly summarised below. The dynamic problems which lead to the algebraic definitions presented next are discussed in [Kar. & Kouv. -2,3], [Jaf. & Kar. -1]. We define [MacF. & Kar. -1], [Kar. -1].

Definition (3.3): Let $S(A,B)$ be a linear system, $N \in \mathbb{R}^{(n-1)xn}$ be a left annihilator of B and let $V \subset X$ be a subspace.

(a) The pencil (sN-NA)V, where V is a basis matrix of V is called a V-restriction pencil (V-r.p.) and the set of strict equivalence invariants of (sN-NA)V will be denoted by $I_V(A,B)$.

(b) The subspace V will be called:

(i) A finite elementary divisor subspace (f.e.d.s.), if $I_V(A,B)$ contains only f.e.d. and possibly zero r.m.i.

(ii) An infinite elementary divisor subspace (i.e.d.s.), if $I_V(A,B)$ contains only i.e.d. and possibly zero r.m.i.

(iii) A column minimal indices subspace (c.m.i.s.), if $I_V(A,B)$ contains only c.m.i. and possibly zero r.m.i.

(iv) A row minimal indices subspace (r.m.i.s.), if $I_V(A,B)$ contains only r.m.i. □

The families of f.e.d.s., i.e.d.s., c.m.i.s., r.m.i.s. will be denoted by $V_f(A,B)$, $V_\infty(A,B)$, $V_\varepsilon(A,B)$, $V_\eta(A,B)$ respectively. The relationships between those families and the families $V(A,B)$, $V_\alpha(A,B)$, $R(A,B)$, $R_\alpha(A,B)$ are defined by the following classification theorem [Juff. & Kar. -1].

Theorem (3.1): Let $S(A,B)$ be a linear system and let $V \subset X$ be a subspace. Then,

(i) $V \in V(A,B)$, iff $\exists V_f \in V_f(A,B)$ and $V_\varepsilon \in V_\varepsilon(A,B)$ such that $V = V_f \oplus V_\varepsilon$

(ii) $V_\varepsilon(A,B) = R(A,B)$.

(iii) $V \in V_\alpha(A,B)$, iff $\exists V_f \in V_f(A,B)$, $V_\infty \in V_\infty(A,B)$ and $V_\varepsilon \in V_\varepsilon(A,B)$ such that $V = V_f \oplus V_\infty \oplus V_\varepsilon$.

(iv) $V \in R_\alpha(A,B)$, iff $\exists V_\infty \in V_\infty(A,B)$, $V_\varepsilon \in V_\varepsilon(A,B)$ such that $V = V_\infty \oplus V_\varepsilon$.

□

Clearly $\mathcal{V}_\infty(A,B) \subset \mathcal{R}_a(A,B)$. The subspaces $\mathcal{V}_\infty \in \mathcal{V}_\infty(A,B)$ and $\mathcal{V}_f \in \mathcal{V}_f(A,B)$ have been called infinite spectrum invariant subspaces, fixed spectrum invariant subspaces respectively [Kar. -1]. Subspaces such as $\mathcal{V}_f$, $\mathcal{V}_\infty$ have been defined by Willems [Will. -1] and they have been referred to as coasting, sliding subspaces correspondingly.

The matrix pencil approach leads to a procedure for the computation of the supremal subspaces $\mathcal{V}^*(\mathcal{K})$, $\mathcal{V}_\alpha{}^*(\mathcal{K})$, $\mathcal{R}^*(\mathcal{K})$ and $\mathcal{R}_\alpha{}^*(\mathcal{K})$ which are contained in a given subspace $\mathcal{K}$, by making use of the Kronecher form of (sN-NA)V [Kar. -1], [Jaff. & Kar. -1].

Corollary (3.1): Let $\mathcal{K}$ be a subspace of $\mathcal{X}$. There always exist subspaces $\mathcal{V}_\eta \in \mathcal{V}_\eta(A,B)$, $\mathcal{V}_f \in \mathcal{V}_f(A,B)$, $\mathcal{V}_\infty \in \mathcal{V}_\infty(A,B)$ and $\mathcal{V}_\varepsilon \in \mathcal{V}_\varepsilon(A,B)$ such that

$$\mathcal{K} = \mathcal{V}_\eta \oplus \mathcal{V}_f \oplus \mathcal{V}_\infty \oplus \mathcal{V}_\varepsilon \tag{3.12}$$

Furthermore, we have that

(i) $\mathcal{V}^*(\mathcal{K}) = \mathcal{V}_f \oplus \mathcal{V}_\varepsilon$, $\mathcal{R}^*(\mathcal{K}) = \mathcal{V}_\varepsilon$

(ii) $\mathcal{V}_\alpha{}^*(\mathcal{K}) = \mathcal{V}_f \oplus \mathcal{V}_\infty \oplus \mathcal{V}_\varepsilon$, $\mathcal{R}_\alpha{}^*(\mathcal{K}) = \mathcal{V}_\infty \oplus \mathcal{V}_\varepsilon$ $\square$

3.3 Generalised autonomous differential systems, and matrix pencils in linear systems theory

The objective of the work presented in this thesis is the study of the algebraic, geometric and dynamic properties of the set of linear, autonomous, first-order differential equations.

$$S(F,G):\quad Fp\underline{x}(t) = G\underline{x}(t),\quad p = \frac{d}{dt},\quad F,G \in \mathbb{R}^{m\times n} \tag{3.13}$$

where $\underline{x}(\cdot): (0^-,\infty) \to \mathbb{R}^n$. Differential systems of the above type are intimately related with the theory of matrix pencils, since the algebraic, geometric and dynamic properties stem from the structure by the associated matrix pencil $sF-\hat{s}G$. In general, $S(F,G)$ systems do not always represent physical dynamical systems, since as we shall see, for a given initial condition, a solution may not be uniquely defined. It is the purpose of this section to show that $S(F,G)$ type systems are intimately related to problems of linear state space theory of the regular, or extended state space type; thus, matrix pencil theory will emerge as the key operator description in the study of linear systems. The formal unification of linear state space theory by means of $S(F,G)$ descriptions and the matrix pencil theory, has been initiated by the work in [MacF. & Kar. -1], [Kar. -1], [Kar. & Kouv. -1], [Jaf. & Kar. -1], [Kar. & MacB. -1], [Kar. & Hay. -1,2], [Apl. -1]; the origins of the "matrix pencil approach", however, go back to the work of Rosenbrock [Ros. -1], and Kalman [Kal. -1].

In justification of a study of (3.13) consider the system

$$S_e(E,A',B',C,D): \begin{aligned} E\,\dot{\underline{x}}(t) &= A'\underline{x}(t)+B'\underline{u}(t) && (3.14a)\\ \underline{y}(t) &= C'\underline{x}(t)+D'\underline{u}(t) && (3.14b)\end{aligned}$$

where $E,A' \in \mathbb{R}^{nxn}$, $B' \in \mathbb{R}^{nxl}$, $C' \in \mathbb{R}^{mxn}$, $D' \in \mathbb{R}^{mxl}$. B' is assumed to have full rank, but E may be singular. If E has full rank $S_e(E,A',B',C',D')$ is reduced to the standard regular state space description

$$S(A,B,C,D): \begin{aligned} \dot{\underline{x}} &= A\,\underline{x}(t)+B\,\underline{u}(t) \quad A = E^{-1}A',\ \ B = E^{-1}B' && (3.15a)\\ \underline{y} &= C\,\underline{x}(t)+D\,\underline{u}(t) && (3.15b)\end{aligned}$$

whereas, if E singular we have the extended state space description $S_e(E,A',B',C',D')$. A number of equivalent descriptions of S_e, or S of the $S(F,G)$ type may be derived by inspection of (3.14), or (3.15) respectively. Thus, by defining the composite vectors $\underline{\zeta}(t)$, $\underline{\xi}(t)$ where $\underline{\zeta}(t)=[\underline{x}(t)^t,\ \underline{u}(t)^t]^t$, $\underline{\xi}(t)=[\underline{x}(t)^t,\ \underline{u}(t)^t,\ \underline{y}(t)^t]^t$ we obtain the following equivalent descriptions of (3.14) [Kar. & Hay. -1,2].

$$S(\Gamma,\Delta,\Theta): \underbrace{\begin{bmatrix} E_n & 0 \\ 0 & 0_\ell \end{bmatrix}}_{\triangleq\,\Gamma} \underbrace{\begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{u}}(t) \end{bmatrix}}_{\triangleq\,\underline{\zeta}(t)} = \underbrace{\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}}_{\triangleq\,\Delta} \begin{bmatrix} \underline{x}(t) \\ \underline{u}(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ -I_m \end{bmatrix}}_{\triangleq\,\Theta} \underline{y}(t) \quad (3.16)$$

or

$$S(\Phi,\Omega): \underset{\triangleq\,\Phi}{\begin{bmatrix} E_\eta & 0 & 0 \\ 0 & 0_\ell & O_m \end{bmatrix}} \underset{\triangleq\,\underline{\dot{\xi}}(t)}{\begin{bmatrix} \underline{\dot{x}}(t) \\ \underline{\dot{u}}(t) \\ \underline{\dot{y}}(t) \end{bmatrix}} = \underset{\triangleq\,\Omega}{\begin{bmatrix} A' & B' & 0 \\ C' & D' & -I_m \end{bmatrix}} \begin{bmatrix} \underline{x}(t) \\ \underline{u}(t) \\ \underline{y}(t) \end{bmatrix} \qquad (3.17)$$

Implicit in the above descriptions is of course the differentiability of the $\underline{u}(t)$, $\underline{y}(t)$; this may be guaranteed by appropriate definition of the space of control inputs. A representation similar in nature to $S(\Gamma,\Delta,\Theta)$ has also appeared in [Appl. -1], [Ber. -1] and has been termed as an implicit <u>description</u>. Such a description is defined by

$$\begin{bmatrix} E_\eta \\ 0 \end{bmatrix} \underline{\dot{x}}(t) = \begin{bmatrix} A' \\ C' \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 & B' \\ -I_m & D' \end{bmatrix} \begin{bmatrix} \underline{y}(t) \\ \underline{u}(t) \end{bmatrix} \qquad (3.18)$$

The above description has emerged as a natural representation from the work on non-oriented models [Apl. -2].

The $S(\Phi,\Omega)$ description is clearly of the $S(F,G)$ type and its importance will be emphasized when the problem of dual systems is addressed later on [Kar. & Hay. -2,3]. The importance of the $S(\Gamma,\Delta,\Theta)$ representation stems from its links to the dynamic characterisation of finite and infinite zeros [MacF. & Kar. -2], [Kar. & Hay. -3], as well as to the theory of multivariable Nyquist and Root locus [MacF. & Kar. -1], [Kar. & Hay. -1]. Thus, by setting $\underline{u}(t)=\underline{0}$ (3.16) yields

$$S(\Gamma,\Delta): \qquad \Gamma\,\underline{\dot{\xi}}(t) = \Delta\,\underline{\xi}(t) \qquad (3.19)$$

which expresses the general output zeroing problem on S_e [Kar. & Kouv. -1], [MacF. & Kar. -2], [Kar. & Hay. -3]. The $S(\Gamma,\Delta)$ system has been defined as the <u>output zeroing differential system</u> and its solutions define the zero structure of S_e.

Variable structure $S(F,G)$ type systems arise in problems such as the Multivariable Nyquist and Root locus ([Post. & MacF. -1], [Kouv. & Sha. -1]);

thus, assume $m=\ell$ and let $\underline{u}(t)=k\underline{y}(t)$, where k represents a scalar output feedback. By substituting in (3.16) and setting $g=k^{-1}$, we obtain the "closed-loop" description

$$S(\Gamma,\Delta(g)): \quad \begin{bmatrix} E_\eta & 0 \\ 0 & 0_\ell \end{bmatrix} \begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{u}}(t) \end{bmatrix} = \begin{bmatrix} A' & B' \\ C' & -gI+D' \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{u}(t) \end{bmatrix} \qquad (3.20),$$

$S(\Gamma,\Delta(g))$ is a variable structure generalised differential system and has been defined as the <u>scalar-gain closed-loop differential system</u> [Kar. & Hay. -1]. The analysis of such systems differes from that of the fixed structure $S(F,G)$ systems; in fact, the fixed structure case is characterised by the pencil pF-G (p=d/dt), whereas the variable structure is characterised by the matrix net

$$P_e(p,g) = p\Gamma + g\Xi - \Delta = p\begin{bmatrix} E_\eta & 0 \\ 0 & 0_\ell \end{bmatrix} + g\begin{bmatrix} 0_\eta & 0 \\ 0 & J_\ell \end{bmatrix} - \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \qquad (3.21)$$

$P_e(p,g)$ has been defined in [MacF. & Kar. -1] as the <u>closed loop system net</u>; for the case where $E=I_\eta$, $P(p,g)$ is related to the characteristic gain-frequency algebraic functions [MacF. & Kar. -1] which in turn define the generalised Nyquist and Root locus ([Post. & MacF. -1], [Smi. -1]). The system $S(\Gamma,\Delta(g))$ may be thought of as the singularly perturbed $S(\Gamma,\Delta)$ system; in fact, if $g\to 0$ ($K\to\infty$, asymptotic root locus case), $\lim_{g\to 0} S(\Gamma,\Delta(g))=S(\Gamma,\Delta)$ which clearly demonstrates the equivalence between the asymptotic root locus and output zeroing problems [Kar. & Hay. -3].

Alternative generalised autonomous descriptions for S_e system problems arise in the matrix pencil characterisation of the various notions of invariant subspaces of the geometric theory ([Kar. -1], [Jaf. & Kar. -1]). Such differential systems are defined next in the more general context of $S_e(E,A',B')$ descriptions. Thus, let us assume $\ell<n$ and let $N' \in \mathbb{R}^{(n-\ell)xn}$, $B'^{+} \in \mathbb{R}^{\ell xn}$ be matrices such that

$$N'B' = 0_{n-\ell,\ell}, \quad B'^{+}B' = I_{\ell} \text{ and } Q = \begin{bmatrix} N' \\ \hline B'^{+} \end{bmatrix} \in \mathbb{R}^{nxn}, \quad |Q| \neq 0 \tag{3.22}$$

where N' is a left annihilator of B' and B'^{+} is a left inverse of B'. By premultiplying (3.14a) by Q, the following equivalent conditions is obtained

$$N'E\,\dot{\underline{x}}(t) = N'A'\underline{x}(t) \tag{3.23a}$$

$$\mathcal{S}_e(N'E,N'A'): \quad \underline{u}(t) = B'^{+}\{E\,\dot{\underline{x}}(t) - A'\underline{x}(t)\} \tag{3.23b}$$

$$\underline{y}(t) = C'\underline{x}(t) + D'\underline{u}(t) \tag{3.23c}$$

Clearly, (3.23a) is of the $\mathcal{S}(F,G)$ type; for any solution of (3.23a) $\underline{x}(t)$, the corresponding $\underline{u}(t)$ that generates $\underline{x}(t)$ and the resulting $\underline{y}(t)$ are given by (3.23b,c). The $\mathcal{S}_e(N'E, N'A')$ description has been termed the <u>input-space restricted state mechanism model</u> [Kar. & Huy. -1]. For $\mathcal{S}(A,B)$ systems equations (3.23a), (3.23b) take the simpler form

$$\mathcal{S}(N,NA): \quad \begin{aligned} N\,\dot{\underline{x}}(t) &= NA\,\underline{x}(t) && (3.24a) \\ \underline{u}(t) &= B^{+}\{\dot{\underline{x}} - A\,\underline{x}(t)\} && (3.24b) \end{aligned}$$

If we now consider the family of systems derived from (A,B) under state feedback i.e. $\mathcal{S}(A{+}BL,B)$: $\dot{\underline{x}}(t)=(A+BL)\underline{x}(t)+B\underline{u}(t)$, then

$$N\,\dot{\underline{x}}(t) = N(A+BL)\,\underline{x}(t) \qquad N\,\dot{\underline{x}}(t) = NA\,\underline{x}(t) \tag{3.25}$$

The above equation demonstrates that $\mathcal{S}(N,NA)$ does not characterise a particular element of $\mathcal{S}(A{+}BL,B)$ but the orbit itself; it is for this reason that $\mathcal{S}(N,NA)$ has been termed as a "feedback free" description [Kar. -1].

Differential systems of $\mathcal{S}(F,G)$ type also arise in the study of solutions of $\mathcal{S}_e$, or $\mathcal{S}$ which are restricted to a given subspace $\mathcal{V}$ of $\mathbb{R}^n$. This problem Provides the basis for the dynamic, geometric and algebraic characterisation of subspaces of $\mathcal{S}_e$, or $\mathcal{S}$. Thus, let $\underline{x}(t) \in \mathcal{V}$ and write $\underline{x}(t)=V\underline{v}(t)$, where V is a basis matrix for $\mathcal{V}$. Then eq (3,23a,b) yield

$$\mathcal{S}_e(N'EV,N'A'V): \quad \begin{aligned} N'EV\,\dot{\underline{v}}(t) &= N'A'V\underline{v}(t), \quad x(t) = V\,\underline{v}(t) && (3.26a) \\ u(t) &= B'^{+}\{E\,\dot{\underline{x}}(t) - A'\,\underline{x}(t)\} && (3.26b) \end{aligned}$$

The system $S_e(N'EV, N'A'V)$ has been termed as the <u>input and V space restricted state mechanism model</u> of S_e. Note that $S_e(N'EV, N'A'V)$ is the key differential system for the developement of an algebrogeometric theory of subspaces of S_e starting from simple dynamic problems; the formulation given above is an extension of the formulation given by Karcanias [Kar. -1] for $S(A,B,C)$ systems and which have led to the matrix pencil characterisation of the geometric concepts ([Kar. -1], [Jaf. & Kar. -1]).

The autonomous differential systems defined above are intimately related with matrix pencils in the variable p=d/dt, or s, the Laplace variable. A summary of the various pencils which are important in linear systems is given in Table (3.1). For the regular state space case, the importance of the various pencils for linear systems is well established; for the extended state space theory although some results, similar in nature to the regular case, are known, but the general theory is not fully developed. Thus, T(p) characterises the pole structure (eigenvalues, eigenspaces). The pencils P(p), Z(p) characterise the zero structure (zeros, elementary output nulling subspaces) [MacF. & Kar. -2], [Kar. & Kav. -1]; the relationships between the invariants of the two pencils are given in [Kar. & MacB. -1]. The pencils C(p), Q(p) define the controllability properties and the pencils R(p), Y(p) characterise the observability properties; the relationships between the invariants of the corresponding pencils is given in [Kar. & MacB. -1]. The system net P(p,g) is instrumental in defining the characteristic frequency and characteristic gain algebraic functions [MacF. & Kar. -1] and thus it is related to the generalised Nyquist and Root locus theory [Post. & MacF. -1]. The pencil Ξ(p) is the key tool in the algebraic characterisation of the various notions of invariant subspaces of the geometric theory [Kar. -1], [Jaf. & Kar. -1].
The theory of invariants and canonical forms in linear systems is based on the theory of strict equivalence of matrix pencils [Kal. -1], [Ros. -1],

Table (3.1): Matrix pencils in linear systems theory

Extended state space	Regular state space	Matrix pencil
$E\dot{\underline{x}} = A'\underline{x}+B'\underline{u}$ $\underline{y} = C'\underline{x}+D'\underline{u}$	$\dot{\underline{x}} = A\underline{x}+B\underline{u}$ $\underline{y} = C\underline{x}+D\underline{u}$	pF-G, p=d/dt
$pE-A' \triangleq T_e(p)$	$pI-A \triangleq T(p)$	Pole pencil
$[pE-A',-B'] \triangleq C_e(p)$	$[pI-A,-B] \triangleq C(p)$	Input-state pencil
$\begin{bmatrix} pE-A' \\ -C' \end{bmatrix} \triangleq R_e(p)$	$\begin{bmatrix} pI-A \\ -C \end{bmatrix} \triangleq R(p)$	State-output pencil
$\begin{bmatrix} pE-A' & -B' \\ -C' & -D' \end{bmatrix} \triangleq P_e(p)$	$\begin{bmatrix} pI-A & -B \\ -C & -D \end{bmatrix} \triangleq P(p)$	Rosenbrock's system matrix pencil
$\begin{bmatrix} pE-A' & -B' & 0 \\ -C' & -D' & -I \end{bmatrix} \triangleq \Theta_e(p)$	$\begin{bmatrix} pI-A & -B & 0 \\ -C & -D & -I \end{bmatrix} \triangleq \Theta(p)$	Input-state-output pencil
$pN'E-N'A' \triangleq Q_e(p)$	$pN-NA \triangleq Q(p)$	Restricted-input-state pencil
$pEM'-A'M' \triangleq Y_e(p)$ M,M': Right annihilators of C',C	$pM-AM \triangleq Y(p)$	Restricted-state-output pencil
$pN'EM'-N'A'M' \triangleq Z_e(p)$ $(D'=0)$	$pNM-NAM \triangleq Z(p)$ $(D=0)$	Zero pencil
$pN'EV-N'A'V \triangleq \Xi_e(p)$	$pNV-NAV \triangleq \Xi(p)$	V-restriction pencil
$\begin{bmatrix} pE-A' & -B' \\ -C' & gI-D' \end{bmatrix} \triangleq P_e(p,g)$ $m=\ell$	$\begin{bmatrix} pI-A & -B \\ -C & gI-D \end{bmatrix} \triangleq P(p,g)$ $m=\ell$	Closed-loop system net

[Thor. -1], [Mor. -1], [Bru. -1], [Kar. & MacB. -1]; the transformation groups associated with the particular canonical forms are defined by the special strict equivalence transformations which reduce a pencil from a given class to a Kronecher canonical form from the same class of pencils.

3.4 Autonomous differential systems, matrix pencils and duality

The generalised autonomous differential system $S(F,G)$ has emerged as a unifying description for the study of properties in linear systems. Their importance has been further emphasized by their role in defining some important notions of duality [Kar. & Hay. -1,3]. The notion of "dual configuration" and of the "dual problem" originates from Projective geometry [Grun. & Wei. -1]; the essence of the "dual systems" defined next, is of similar nature to that of Projective geometry and is defined by the properties of strict equivalence invariants of homogeneous matrix pencils. The importance of the dual differential systems stems from that a "Principle of Duality" similar to that of Projective geometry may be stated for autonomous differential system; using this principle, if a proposition is true on one system, the dual proposition is true for the dual system.

For a pair $(F,G) \in \mathbb{R}^{mxn} x \mathbb{R}^{mxn}$, we may always associate with it the differential systems [Kar. & Hay. -1]

$$S(F,G): \quad F\,p\,\underline{x}(t) = G\,\underline{x}(t) \tag{3.27}$$

$$\tilde{S}(F,G): \quad F^t p\,\underline{\tilde{x}}(t) = G^t\underline{\tilde{x}}(t) \tag{3.28}$$

$$\hat{S}'(F,G): \quad F\,\underline{\hat{x}}(t) = G\,p\,\underline{\hat{x}}(t) \Leftrightarrow F\,p^{-1}\underline{\hat{x}}(t) = G\,\underline{\hat{x}}(t) \tag{3.29}$$

$S(F,G)$ will be referred to as the prime system, $\tilde{S}(F,G)$ as the transposed dual and $\hat{S}(F,G)$ as the proper dual system, or simply dual. The duality between $S(F,G)$ and $\tilde{S}(F,G)$ has been termed as transposed duality and the duality between $S(F,G)$ and $\hat{S}(F,G)$ as differentiator-integrator duality, since p is the derivative operator and p^{-1} the integrator operator. The following diagram summarises the relationships between the three systems

$$\begin{tikzcd}
 & \mathcal{S}(F,G) \arrow[dl, leftrightarrow, "\text{TRANSPOSED DUALITY}"'] \arrow[dr, leftrightarrow, "\text{INTEGRATOR-DIFFER. DUALITY}"] & \\
\tilde{\mathcal{S}}(F,G) \arrow[rr, leftrightarrow, "\text{TRANSPOSED AND INTEGRATOR-}", "\text{DIFFERENTIATOR DUALITY}"'] & & \hat{\mathcal{S}}(F,G)
\end{tikzcd}$$

Figure (3.1)

Given that with $\mathcal{S}(F,G)$, $\tilde{\mathcal{S}}(F,G)$, $\hat{\mathcal{S}}(F,G)$ we may associate the pencils $pF-G$, pF^t-G^t, $F-pG$ respectively, it is clear that the essence of the various types of duality depends on the relationships between the associated pencils. In the following we shall assume that the indeterminates in the pencils are the Laplace variables; infact, by taching Laplace transforms of (3.27), (3.28), (3.29) we have

$$\mathcal{S}(F,G): \rightarrow (sF-G)\,x(s) = F\,\underline{x}(0^-) \tag{3.30}$$

$$\tilde{\mathcal{S}}(F,G): \rightarrow (\tilde{s}F^t-G^t)\,\underline{\tilde{x}}(\tilde{s}) = F^t\,\underline{\tilde{x}}(0^-) \tag{3.31}$$

$$\hat{\mathcal{S}}(F,G): \rightarrow (F-\hat{s}G)\,\underline{\hat{x}}(\hat{s}) = G\,\underline{\hat{x}}(0^-) \tag{3.32}$$

where $\underline{x}(0^-)$, $\underline{\tilde{x}}(0^-)$, $\underline{\hat{x}}(0^-)$ denote the initial conditions at $t=0^-$ of the corresponding vectors, s, $\tilde{s}$, $\hat{s}$ denote the Laplace variable and $x(s)$, $\tilde{x}(s)$, $\hat{x}(s)$ the Laplace transforms of $x(t)$, $\tilde{x}(t)$, $\hat{x}(t)$ respectively [Doe. -1]. The study of duality between $\mathcal{S}(F,G)$, $\tilde{\mathcal{S}}(F,G)$, $\hat{\mathcal{S}}(F,G)$, is reduced to an investigation of the links between the pencils $sF-G$, $\tilde{s}F^t-G^t$, $F-\hat{s}G$; it is clear that the homogeneous pencil $sF-\hat{s}G$ is the appropriate tool for the establishment of the duality relationships.

The notions of duality defined above may be qualified algebraically in terms of relationships between the strict equivalence invariants of the associated pencils. These relationships are summarised below [Kar. & Hay. -1, 2].

<u>Theorem (3.2)</u>: The strict equivalence invariants of the pencils sF-G, F-$\hat{s}$G defined on a general pair $(F,G) \in \mathbb{R}^{mxn} x \mathbb{R}^{mxn}$ are related as follows:

(i) A z.e.d. s^p of sF-G is an i.e.d. of F-$\hat{s}$G and vice-versa.

(ii) An i.e.d. $\hat{s}^q$ of $\hat{s}$F-G is a z.e.d. of F-$\hat{s}$G and vice-versa.

(iii) A n.-z.f.e.d. $(s-a)^\tau$ of sF-G defines a n.-z.f.e.d. $(\hat{s}-1/a)^\tau$ of F-$\hat{s}$G and vice versa.

(iv) The sets of c.m.i. and r.m.i. of sF-G and F-$\hat{s}$G are equal.

□

Clearly, the dominant characteristic in this type of duality is the "inversion" of frequency which is defined by the dual role of the different types of e.d.; since this duality is dominated by the relationships between the set of e.d., it has been called as <u>elementary divisors type duality</u> ([Kar. & Hay. -1,2]). It is worth pointing out that for singular pencils, not only the degrees but also the real spaces associated with polynomial vectors in N_r(sF-G), N_r(F-$\hat{s}$G) (N_ℓ(sF-G), N_ℓ(F-$\hat{s}$G)) are related. Thus, we have [Kar. & Hay. -2]:

<u>Corollary (3.2)</u>: Let $\underline{x}(s) = \underline{x}_o + \underline{x}_1 s + \ldots + \underline{x}_n s^k \in \mathbb{R}^n[s]$ be such that $(sF-G)\,\underline{x}(s) = \underline{0}$. Then $\underline{x}(s) = \underline{x}_o s^k + \underline{x}_1 s^{k-1} + \ldots + \underline{x}_k \in \mathbb{R}^n[s]$ also satisfies $(F-\hat{s}G)\,\hat{\underline{x}}(\hat{s}) = \underline{0}$ and vice versa.

□

Clearly, the space $\mathcal{R}_x = sp\{\underline{x}_o, \underline{x}_1, \ldots, \underline{x}_k\}$ characterises both vectors $\underline{x}(s)$, $\hat{\underline{x}}(\hat{s})$ and thus it is an invariant under this duality. A similar result may be stated for the generalised eigenspaces associated with the various types of e.d. (finite and infinite). For the duality between sF-G and $\hat{s}F^t-G^t$ we have [Kar. & Hay. -1].

Theorem (3.3): The strict equivalence invariants of the pencils $sF-G$, $\tilde{s}F^t-G^t$ defined on a general pair $(F,G) \in \mathbb{R}^{mxn} \times \mathbb{R}^{mxn}$ are related as follows:

(i) The sets of z.e.d., n.-z.f.e.d., i.e.d. of $sF-G$ and $\tilde{s}F^t-G^t$ are equal (modulo change of variables s to $\tilde{s}$).

(ii) The set of c.m.i. of $sF-G$ is equal to the set of r.m.i. of $\tilde{s}F^t-G^t$.

(iii) The set of r.m.i. of $sF-G$ is equal to the set of c.m.i. of $\tilde{s}F^t-G^t$.

$\square$

It is because of the above property that such duality has been referred to as minimal index type duality. By combining the above two theorems we may express the duality between $sF-G$ and $F^t-\bar{s}G^t$ as follows:

Theorem (3.4): The strict equivalence invariant of the pencils $sF-G$, $F^t-\bar{s}G^t$ defined on a general pair $(F,G) \in \mathbb{R}^{mxn} \times \mathbb{R}^{mxn}$ are related as follows:

(i) A z.e.d. s^p of $sF-G$ is an i.e.d. of $F^t-\bar{s}G^t$ and vice versa.

(ii) An i.e.d. $\bar{s}^q$ of $sF-G$ is a z.e.d. of $F^t-\bar{s}G^t$ and vice versa.

(iii) A n.-z.f.e.d. $(s-a)^\tau$ of $sF-G$ defines a n.-z.f.e.d. $(\bar{s}-1/a)^\tau$ of $F^t-\bar{s}G^t$ and vice versa.

(iv) The set of c.m.i. of $sF-G$ is equal to the set of r.m.i. of $F^t-\bar{s}G^t$.

(v) The set of r.m.i. of $sF-G$ is equal to the set of c.m.i. of $F^t-\bar{s}G^t$.

$\square$

The differential system defined on $(F,G) \in \mathbb{R}^{mxn} \times \mathbb{R}^{mxn}$ by

$$\bar{S}(F,G): \quad F^t \bar{\underline{x}}(t) = G^t p \bar{\underline{x}}(t) \tag{3.33}$$

will be called the completely dual system of $S(F,G)$ and the duality between $S(F,G)$, $\bar{S}(F,G)$ will be referred to as complete duality, or elementary divisor-minimal index duality. The duality notions discussed here may also be characterised dynamically (in terms of relationships between the basic solutions associated with each type of invariants); such a property has been used in [Kar. & Hay. -3] for the characterisation of infinite zeros of state space models in terms of standard frequency transmission blocking problems [MacF. & Kar. -2] on appropriately defined dual system. It is worth noting that in this case, a uniquely defined impulsive solution (associated with i.e.d.) becomes the dual of a uniquely defined solution involving step, rump, etc. functions (associated with a z.e.d.). The various aspects of the duality notions will be further clarified in the following chapters.

3.5 Conclusions

The differential system $S(F,G)$ has emerged as the unifying description for the study of properties and problems of regular and extended state space linear systems. The importance of $S(F,G)$ descriptions is due to its intimate links to a rather simple in form, but rich in structure linear operator, the matrix pencil sF-G. The classical theory of matrix pencils deals with the algebraic structure of sF-G under strict equivalence. Almost all new developements in the theory of matrix pencils have been in the area of numerical analysis [Wilk. -1,2,3], [Stew. -1,2], [Van Dor. -1,3] etc.; however, until recently [Kar. & Hay. -1,2], [Ber. -1] the other aspects of the matrix pencil theory have not taken any significant attention. From the linear systems viewpoint, there is a need for further developement of the classical theory especially in the direction of the geometric aspects of the theory. One of the major aims of the present work is to contribute in the enrichment of the classical theory of matrix pencils by exploring the geometric, number theoretic and more general invariant theory aspects; by transferring the new results on the differential system description $S(F,G)$ the dynamic and "relativistic" aspects of the theory also emerge. The intimate link of $S(F,G)$ descriptions with linear systems theory, indicates that matrix pencil theory, enriched with new results, may provide a general theory for linear systems based on the structure and properties of a single linear operator.

The importance of Segre' theory in the study of the standard eigenvalue-eigenvector problem motivates the study of extending the results from the context of the simple pencil sI-A to the general case sF-G. A geometric algebraic and dynamic theory of subspaces of the domain of (F,G) is developed. The characterisation of invariants from the number theoretic viewpoint provides the means for novel calculations of canonical forms of matrix pencils. The duality of $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ systems motivates the study of algebraic and geometric invariants of matrix pencils under Bilinear strict equivalence; it is through this study, that a "space-frequency relativistic" classification of properties of $\mathcal{S}(F,G)$ descriptions, and thus of linear systems, may be given.

CHAPTER 4:

Number Theoretic and Geometric aspects of the strict equivalence invariants of Regular Pencils

CHAPTER 4: NUMBER THEORETIC AND GEOMETRIC ASPECTS OF THE STRICT EQUIVALENCE INVARIANTS OF REGULAR PENCILS

4.1 Introduction

The aim of this chapter is to provide a detailed study of the number theoretic and geometric aspects of the strict equivalence invariants of a regular homogeneous pencil $sF-\hat{s}G$. The main idea running throughout this chapter is the generalisation of the number theoretic and geometric aspects of the standard eigenvalue-eigenvector problem associated with the regular pencil $sI-A$, $A \in \mathbb{R}^{n\times n}$, to the case of the generalised eigenvalue problem associated with the regular pencil $sF-G$, $F,G \in \mathbb{R}^{n\times n}$, with F not necessarily invertible.

The e.d. structure of the pencil $sI-A$ may be characterised from the number theoretic point of view by the Segre and Weyr characteristics of each distinct eigenvalue of A; the computation of the Jordan form is then reduced to a problem of constructing a Ferrer's diagram [Tur. & Ait - 1]. An essential part in the computation of a similarity transformation, which reduces A to its Jordan normal form is the notion of the generalised null space of A, defined for each distinct eigenvalue of A [Dor. - 1]; in fact the selection of linearly independent chains of generalised eigenvectors, characterising each Jordan block corresponding to an eigenvalue λ_i, makes use of the properties of the generalised nullspace of A. The Segre, Weyr characteristics provide the number theoretic aspects of $sI-A$, whereas the properties of the generalised nullspaces (for the various distinct eigenvalues) provide the basis for the study of the geometric properties of $sI-A$, as they are expressed by the chains of generalised eigenvectors.

With every root of an e.d. of $sF-\hat{s}G$, finite or infinite, a subspace spanned by generalised eigenvectors is associated. This, provides the means for the characterisation of the set of e.d. at $s = \alpha \in \mathbb{C} \cup \{\infty\}$, as number invariants, defined by the nullspace properties of the α-Toeplitz matrices $P^k_\alpha(F,G)$, $K=1,2,\dots$, which are completely specified by α and the pair (F,G). The notions of Segre, Weyr characteristics and the Ferrer's diagram are extended to the case of $sF-\hat{s}G$. An alternative way for computing the Segre characteristic at $s=\alpha$ (different than that based on the Ferrer's diagram) is given; this new approach is based on properties of sequences defined by α and (F,G), which satisfy the piecewise arithmetic progression property.

The study of the geometry of the set of e.d. at $s=\alpha$ is based on the properties of $\mathcal{N}_r\{P^k_\alpha(F,G)\} = \mathcal{N}^k_\alpha$; the key tool in this study is the notion of a nested basis matrix N^k_α of $\mathcal{N}^k_\alpha$. The study of properties of the nested bases matrices of $\mathcal{N}^k_\alpha$ suggests the way we can select maximal independent chains characterising each e.d. at $s=\alpha$ and it also provides a third method for computing the Segre characteristic of (F,G) at $s=\alpha$. The results presented here are generalisations and extensions of well-known facts for $sI-A$ to the more general case of $sF-\hat{s}G$. Such a study provides also the means for an alternative method of constructing the Weierstrass form of a regular pencil.

4.2 The α-Toeplitz matrices and the k-th generalised nullspace of (F,G): preliminary results and definitions

In this section, a number of preliminate results are derived which lead to the definition of the α-Toeplitz matrices associated with the pair (F,G) and to the definition of the k-th generalised nullspace of (F,G) at $s=\alpha$. For the n×n regular pencil $sF-\hat{s}G$, the sets of e.d. corresponding to the same frequency (represented by an ordered pair) shall be denoted by

$$\begin{aligned}&\mathcal{D}(1,0) \triangleq \{s^{p_i}, i\in\underset{\sim}{\mu}, p_1\le\dots\le p_\mu\}, \mathcal{D}(0,1) \triangleq \{\hat{s}^{q_i}, i\in\underset{\sim}{\nu}, q_1\le\dots\le q_\nu\}\\ &\mathcal{D}(1,\alpha) \triangleq \{(s-\alpha\hat{s})^{d_i}, i\in\underset{\sim}{\tau}, \alpha\in\mathbb{C}-\{0\}, d_1\le\dots\le d_\tau\} \quad \text{or} \qquad (4.1)\\ &\mathcal{D}(\hat{\alpha},1) \triangleq \{(\hat{\alpha}s-\hat{s})^{d_i}, i\in\underset{\sim}{\tau}, \hat{\alpha}=1/\alpha\in\mathbb{C}-\{0\}, d_1\le\dots\le d_\tau\}\end{aligned}$$

In the case of sF-G the sets $\mathcal{D}(1,0),\mathcal{D}(0,1),\mathcal{D}(1,\alpha)$ will be denoted simply by $\mathcal{D}(0),\mathcal{D}(\infty),\mathcal{D}(\alpha)$ and represent the sets of zero, infinite and α e.d. respectively. For the case of $F-\hat{s}G$, the $\mathcal{D}(1,0),\mathcal{D}(0,1),\mathcal{D}(\hat{\alpha},1)$ will be denoted simply by $\hat{\mathcal{D}}(\infty),\hat{\mathcal{D}}(0),\hat{\mathcal{D}}(\hat{\alpha})$ correspondingly and represent the sets of infinite, zero and $1/\alpha$ e.d. respectively.

The sets of e.d. $\mathcal{D}(1,0),\mathcal{D}(0,1),\mathcal{D}(1,\alpha)$ define the Weierstrass form $sF_w-\hat{s}G_w$ of sF-sG described by

$$sF_w-\hat{s}G_w = \text{diag}\{\ldots;D_p(s,\hat{s});\ldots;J_d^{\alpha}(s,\hat{s});\ldots;\hat{D}_q(s,\hat{s});\ldots\} \tag{4.2}$$

where the blocks $D_p(s,\hat{s}),J_d^{\alpha}(s,\hat{s}),\hat{D}_q(s,\hat{s})$ are defined by

$$\begin{array}{lrcll} & D_p(s,\hat{s}) & = & sI_p-\hat{s}H_p & \rightarrow\ (s^p) \\ & \hat{D}_q(s,\hat{s}) & = & sH_q-\hat{s}I_q & \rightarrow\ (\hat{s}^q) \\ & J_d^{\alpha}(s,\hat{s}) & = & sI_d-\hat{s}J_d(\alpha) & \rightarrow\ (s-\alpha\hat{s})^d,\ \alpha\neq 0 \\ \underline{\text{OR}} & J_d^{\hat{\alpha}}(s,\hat{s}) & = & sJ_d(\hat{\alpha})-\hat{s}I_d & \rightarrow\ (\hat{\alpha}s-\hat{s})^d,\ \hat{\alpha}=1/\alpha \end{array} \tag{4.3}$$

Note that (4.2) is the Complex Weierstrass form, since the set of e.d. is considered over $\mathbb{C}$. The pencil $sF-\hat{s}G$ is reduced to $sF_w-\hat{s}G_w$ by complex transformations (R,Q), where $R,Q\in\mathbb{C}^{n\times n}$ and $|R|,|Q|\neq 0$.

The elementary divisors of $sF-\hat{s}G$ (finite and infinite) may be associated with maximal length linearly independent vector chains, as it is shown next. The case of finite, non-zero e.d. is considered first and then the results are extended to the 0 and ∞ e.d. The existence of the vector chains is established by the following result.

<u>Proposition (4.1)</u>: Let $sF-\hat{s}G$ be an n×n regular pencil and let $\{(s-\alpha)^{d_i}, i\in\underset{\sim}{r}\}$ be the set of α-e.d. of sF-G and $\{(\hat{s}-\hat{\alpha})^{d_i'}, i\in\underset{\sim}{r}'\}$ be the set of $\hat{\alpha}$-e.d. of $F-\hat{s}G$.

(i) For every $(s-\alpha)^{d_i}$ e.d. there exists a maximal chain of linearly independent vectors $\{\underline{x}_1^i,\underline{x}_2^i,\dots,x_{d_i}^i\}$ such that

$$G\underline{x}_j^i = \alpha F\underline{x}_j^i+F\underline{x}_{j-1}^i,\ j\in\underset{\sim}{d}_i,\ \underline{x}_o^i=\underline{0} \tag{4.4}$$

and with the further property that the set of vectors $\{\underline{x}_1^1,\dots,\underline{x}_{d_1}^1;\dots;$ $\underline{x}_1^\tau,\dots,\underline{x}_{d_\tau}^\tau\}$ is linearly independent.

(ii) For every $(\hat{s}-\hat{\alpha})^{d_i'}$ e.d. there exists a maximal chain of linearly independent vectors $\{\underline{\hat{x}}_1^i,\underline{\hat{x}}_2^i,\dots,\underline{\hat{x}}_{d_i}^i\}$ such that

$$F\underline{\hat{x}}_j^i = \hat{\alpha}G\underline{\hat{x}}_j^i+G\underline{\hat{x}}_{j-1}^i,\ j\in\underset{\sim}{d}_i',\ \underline{\hat{x}}_o^i=0 \tag{4.5}$$

and with the further property that the set of vectors $\{\underline{\hat{x}}_1^1,\dots,\underline{\hat{x}}_{d_1'}^1;\dots;$ $\underline{\hat{x}}_1^{\tau'},\dots,\underline{\hat{x}}_{d_\tau'}^{\tau'}\}$ is linearly independent.

Proof

(i) Let (R,Q) be a pair of matrices reducing sF-G to its Weierstrass form (4.2). Then, all blocks in αF_w-G_w are nonsingular apart from those $J_{d_i}^{\alpha_i}(s)$ for which $\alpha=\alpha_i$. For, every $(s-\alpha)^{d_i}$ e.d. the corresponding block becomes

$$-\alpha I_{d_i}+J_{d_i}(\alpha) = H_{d_i} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \vdots & & & & 1 \\ 0 & \dots & & \dots & 0 \end{bmatrix}$$

By choosing the set of standard basis vectors $\{\underline{e}_1,\dots,\underline{e}_{d_i}\}$, $\underline{e}_i\in\mathbb{R}^{d_i}$, then it is clear that

$$J_{d_i}(\alpha)\underline{e}_j = \alpha I_{d_i}\underline{e}_j+I_{d_i}\underline{e}_{j-1},\ j\in\underset{\sim}{d}_i,\ \underline{e}_o=0 \tag{4.6}$$

The set of vectors $\{\underline{e}_j, j\in d_i\}$ is clearly linearly independent and the chain is maximal since any attempt to generate more independent vectors using the (4.6) algorithm fails. Consider now the vectors $\{\underline{e}_1^i,\dots,\underline{e}_{d_i}^i\}$ of $\mathbb{R}^n$ which have all their coordinates zero apart from the coordinates of the vectors $\{\underline{e}_1,\dots,\underline{e}_{d_i}\}$ at the position of the $-sI_{d_i}+J_{d_i}(\alpha)$ block, i.e.

$$\underline{e}_j^i = [0\ldots0,\underline{e}_j^{\,t},0\ldots0]^t$$

Then

$$G_w\underline{e}_j^i = \alpha F_w\underline{e}_j^i+F_w\underline{e}_{j-1}^i,\ j\in\underset{\sim}{d}_i,\ \underline{e}_o^i=\underline{0}$$

By premultiplying the above by R and using the fact that $QQ^{-1}=Q\hat{Q}=I_n$ we have

$$RG_wQ\hat{Q}\underline{e}_j^i = \alpha RF_wQ\hat{Q}\underline{e}_j^i+RF_wQ\hat{Q}\underline{e}_{j-1}^i,\ j\in\underset{\sim}{d}_i,\ \underline{e}_o^i=0$$

or

$$G\{\hat{Q}\underline{e}_j^i\} = \alpha F\{\hat{Q}\underline{e}_j^i\}+F\{\hat{Q}\underline{e}_{j-1}^i\},\ j\in\underset{\sim}{d}_i$$

By setting $\underline{x}_j^i=\hat{Q}\underline{e}_j^i$, $j\in\underset{\sim}{d}_i$, eqn(4.4) is established. Given that the set of vectors $\{\underline{e}_1^1,\ldots,\underline{e}_{d_1}^1;\ldots;\underline{e}_1^\tau,\ldots,\underline{e}_{d_\tau}^\tau\}$ are independent, the set of vectors $\{\underline{x}_1^1,\ldots,\underline{x}_{d_1}^1;\ldots;\underline{x}_1^\tau,\ldots,\underline{x}_{d_\tau}^\tau\}$ is also independent since $\hat{Q}=Q^{-1}$ has full rank. The proof of part (ii) is similar. □

The existence of e.d. guarantees the existence of maximal and linearly independent chains of vectors which satisfy eqns(4.4) for sF-G, or (4.5) for $F-\hat{s}G$. The inverse problem is considered next.

<u>Proposition (4.2)</u>: Let $sF-\hat{s}G$ be an n×n regular pencil.

(i) If there exists a maximal chain of linearly independent vectors $\{\underline{x}_1,\ldots,\underline{x}_d\}$ such that

$$G\underline{x}_j = \alpha F\underline{x}_j+F\underline{x}_{j-1},\ j\in\underset{\sim}{d},\ \underline{x}_o=\underline{0} \tag{4.7}$$

then sF-G has an e.d. $(s-\alpha)^d$.

(ii) If there exists a maximal chain of linearly independent vectors $\{\underline{\hat{x}}_1,\ldots,\underline{\hat{x}}_{d'}\}$ such that

$$F\underline{\hat{x}}_j = \hat{\alpha}G\underline{\hat{x}}_j+G\underline{\hat{x}}_{j-1},\ j\in\underset{\sim}{d}',\ \underline{\hat{x}}_o=\underline{0} \tag{4.8}$$

then $F-\hat{s}G$ has an e.d. $(\hat{s}-\hat{\alpha})^{d'}$.

Proof

(i) From (4.7) we have that

$$(G-\alpha F)\underline{x}_j = F\underline{x}_{j-1},\ j\in\underset{\sim}{d},\ \underline{x}_o=0 \tag{4.9}$$

and thus for j=1 we have that $(G-\alpha F)\underline{x}_1=0$; this implies that sF-G loses rank at s=α and thus sF-G has an e.d. at s=α. We have to show that the degree of such an e.d. is d. Let (R,Q) be a pair of transformations such that $R(sF-G)Q=sF_w-G_w$. Premultiply (4.9) by R and define the coordinate transformation $\underline{x}_j=Q\underline{\tilde{x}}_j$, $j\in\underset{\sim}{d}$. Then, (4.9) yields

$$(G_w-\alpha F_w)\underline{\tilde{x}}_j = F_w\underline{\tilde{x}}_{j-1},\ j\in\underset{\sim}{d},\ \underline{\tilde{x}}_o=0 \tag{4.10}$$

Assume that all blocks in G_w-sF_w associated with $\{(s-\alpha)^{d_i}, i\in\underset{\sim}{\tau}\}$ e.d. are the last blocks in the Weierstrass form. Then

$$G_w-\alpha F_w = \text{diag}\{\ldots;J_{\tau_i}(\beta_i)-\alpha I_{\tau_i};\ldots;I_{q_i}-\alpha H_{q_i};\ldots;J_{d_1}(\alpha)-\alpha I_{d_1};\ldots \\ \ldots;J_{d_\tau}(\alpha)-\alpha I_{d_\tau}\} \tag{4.11}$$

Clearly all blocks $J_{\tau_i}(\beta_i)-\alpha I_{\tau_i}$, $I_{q_i}-\alpha H_{q_i}$ associated with finite e.d. $(s-\beta_i)^{\tau_i}$, $\beta_i\neq\alpha$ and infinite e.d. $\hat{s}^{q_i}$ respectively are nonsingular; however, the blocks $J_{d_i}(\alpha)-\alpha I_{d_i}=H_{d_i}$ for all $i\in\underset{\sim}{\tau}$ and thus they are singular. Let us now partition the vectors $\underline{\tilde{x}}_j$ according to the partitioning of $G_w-\alpha F_w$ as in (4.11), i.e.

$$\underline{\tilde{x}}_j = [\ldots;\underline{\tilde{x}}^{j\ t}_{\tau_i};\ldots;\underline{\tilde{x}}^{j\ t}_{q_i};\ldots;\underline{\tilde{x}}^{j\ t}_{\alpha}]^t \tag{4.12}$$

where $\underline{\tilde{x}}^j_{\tau_i}$, $\underline{\tilde{x}}^j_{q_i}$ are the subvectors of $\underline{\tilde{x}}_j$ corresponding to the nonsingular blocks $J_{\tau_i}(\beta_i)-\alpha I_{\tau_i}$, $I_{q_i}-\alpha H_{q_i}$ respectively and $\underline{\tilde{x}}^j_{\alpha}$ is the subvector corresponding to the singular matrix $\text{diag}\{J_{d_1}(\alpha)-\alpha I_{d_1};\ldots;J_{d_\tau}(\alpha)-\alpha I_{d_\tau}\}=$ $=\text{diag}\{H_{d_1};\ldots;H_{d_\tau}\}$. Conditions (4.10) are then reduced to the following equivalent set of conditions

$$\{J_{\tau_i}(\beta_i)-\alpha I_{\tau_i}\}\underline{\tilde{x}}^j_{\tau_i} = I_{\tau_i}\tilde{x}^{j-1}_{\tau_i},\ j\in\underset{\sim}{d},\ \underline{\tilde{x}}^o_{\tau_i}=\underline{0} \tag{4.13}$$

$$\{I_{q_i}-\alpha H_{q_i}\}\underline{\tilde{x}}^j_{q_i} = H_{q_i}\underline{\tilde{x}}^{j-1}_{q_i},\ j\in\underset{\sim}{d},\ \underline{\tilde{x}}^o_{q_i}=\underline{0} \tag{4.14}$$

and

$$\mathrm{diag}\{H_{d_1};\dots;H_{d_\tau}\}\underline{\tilde{x}}^j_\alpha = \underline{\tilde{x}}^{j-1}_\alpha,\ j\in\underset{\sim}{d},\ \underline{\tilde{x}}^o_\alpha=\underline{0} \tag{4.15}$$

For j=1, conditions (4.13) and (4.14) have as only possible solution $\underline{\tilde{x}}^1_{\tau_i}=\underline{0}$, $\underline{\tilde{x}}^1_{q_i}=\underline{0}$ respectively since the coefficient matrices are nonsingular; thus all vectors for j>1 in (4.13) and (4.14) are zero and (4.12) yields

$$\underline{\tilde{x}}_j = [0\dots0;\underline{\tilde{x}}^j_\alpha]^t \tag{4.16}$$

For j=1, eqn(4.15) yields that

$$\underline{\tilde{x}}^1_\alpha = [\underbrace{c^1_1,0\dots0}_{\longleftarrow d_1 \longrightarrow};\underbrace{c^2_1,0\dots0}_{\longleftarrow d_2 \longrightarrow};\dots,\underbrace{c^\tau_1,0\dots0}_{\longleftarrow d_\tau \longrightarrow}]^t \tag{4.17a}$$

where $c^j_1\in\mathbb{C}$, $j\in\underset{\sim}{\tau}$ are constants. It is clear that the number of vector chains which may be generated is equal to the number of H_{d_i} blocks, or equal to the number of elementary divisors.

Let us now assume that $d_1=\dots=d_{\rho_1}=f_1<d_{\rho_1+1}=\dots=d_{\rho_2}=f_2<\dots<d_{\rho_{\nu-1}+1}=\dots=d_\tau=f_\nu$ be the ordered set of degrees of the e.d. at s=α and let us define the vectors

$$\underline{\tilde{e}}^i_1 = [\underbrace{0\dots0}_{\leftarrow d_1 \rightarrow};0\dots0;\underbrace{1,0,\dots,0}_{\leftarrow d_i \rightarrow};0\dots0;\underbrace{0\dots0}_{\leftarrow d_\tau \rightarrow}]^t \tag{4.17b}$$

and the subspaces $\mathcal{V}_1=\mathrm{span}\{\underline{\tilde{e}}^1_1,\dots,\underline{\tilde{e}}^{\rho_1}_1\}$, $\mathcal{V}_2=\mathrm{span}\{\underline{\tilde{e}}^{\rho_1+1}_1,\dots,\underline{\tilde{e}}^{\rho_2}_1\},\dots,$ $\mathcal{V}_\nu=\mathrm{span}\{\underline{\tilde{e}}^{\rho_{\nu-1}+1}_1,\dots,\underline{\tilde{e}}^\tau_1\}$. Furthermore, let us denote by $\mathcal{V}_o=\mathcal{V}_1\oplus\dots\oplus\mathcal{V}_i\oplus\dots\oplus\mathcal{V}_\nu$ and by $\mathcal{V}_{1,2,\dots,i)}=\mathcal{V}_{i+1}\oplus\dots\oplus\mathcal{V}_\nu$.

For all $j\le f_1$, eqn(4.15) generates vectors $\underline{\tilde{x}}^j_\alpha$, which may be readily shown to have the following general form

$$\underline{\tilde{x}}^j_\alpha = [\underbrace{c^1_j,\dots,c^1_1,0\dots0}_{\leftarrow\ d_1\ \rightarrow};\dots;\underbrace{c^i_j,\dots,c^i_1,0\dots0}_{\leftarrow\ d_i\ \rightarrow};\dots;\underbrace{c^\tau_j,\dots,c^\tau_1,0\dots0}_{\leftarrow\ d_\tau\ \rightarrow}]^t \tag{4.17c}$$

If $\tilde{\underline{x}}_{\alpha}^{1} \in V_{o}$, but $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1)}$ ($\tilde{\underline{x}}_{\alpha}^{1}$ has a nonzero projection on at least one of the vectors $\{\tilde{\underline{e}}_{1}^{1},\dots,\tilde{\underline{e}}_{1}^{\rho_1}\}$), then by using (4.15) for $j=f_1+1$, it may be readily shown that $c_1^1=\dots=c_1^{\rho_1}=0$ which contradicts the assumption that $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1)}$. Thus, if $\tilde{\underline{x}}_{\alpha}^{1} \notin V_1$ the maximal length of the chain is f_1; note that any chain of length $j<f_1$ is not maximal, since it may be completed by selecting linearly independent vectors up to $j=f_1$, as it is indicated by (4.17c). Therefore, if $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1)}$ the length of the maximal chain is f_1.

Assume next that $\tilde{\underline{x}}_{\alpha}^{1} \in V_{1)}$, but $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1,2)}$ ($\tilde{\underline{x}}_{\alpha}^{1}$ has no projection on $\{\tilde{\underline{e}}_{1}^{1},\dots,\tilde{\underline{e}}_{1}^{\rho_1}\}$, and has at least a nonzero projection on $\{\tilde{\underline{e}}_{1}^{\rho_1+1},\dots,\tilde{\underline{e}}_{1}^{\rho_2}\}$). For all $j\leq f_2$, the vectors generated by (4.15) have the general form

$$\tilde{\underline{x}}_{\alpha}^{j} = [0,\dots,0;\underbrace{c_j^{\rho_1+1},\dots,c_1^{\rho_1+1},0\dots0}_{d_{\rho_i+1}};\dots;\underbrace{c_j^{\tau},\dots,c_1^{\tau},0\dots0}_{d_{\tau}}]^t \tag{4.17d}$$

If the chain has length greater than f_2, then for $j=f_2+1$ it may be readily shown that $c_1^{\rho_1+1}=\dots=c_1^{\rho_2}=0$, which contradicts the assumption that $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1,2)}$ by $\tilde{\underline{x}}_{\alpha}^{1} \in V_{1)}$; thus, in this case the maximal length is f_2. If the chain has length less than f_2, then it cannot be maximal, since (4.17d) may generate more linearly independent vectors up to f_2 in number.

Using similar arguments it may be shown that if $\tilde{\underline{x}}_{\alpha}^{1} \in V_{1,\dots,i)}$ but $\tilde{\underline{x}}_{\alpha}^{1} \notin V_{1,\dots,i,i+1)}$, then the chain has length f_{i+1} and thus the e.d. has degree f_{i+1}. The proof of part (ii) is similar. □

An obvious generalization of the above result is the following proposition.

Proposition (4.3): Let $sF-\hat{s}G$ be an n×n regular pencil.

(i) If there exist maximal chains of linearly independent vectors $\{\underline{x}_1^i,\dots,\underline{x}_{d_i}^i\}$, $i\in\underset{\sim}{\rho}$, such that

$$G\underline{x}_j^i = \alpha F\underline{x}_j^i + F\underline{x}_{j-1}^i, \ j\in\underset{\sim}{d}_i, \ \underline{x}_o^i=0 \tag{4.18}$$

with the vector set $\{\underline{x}_1^1,\dots,\underline{x}_{d_1}^1;\dots;\underline{x}_1^{\rho},\dots,\underline{x}_{d_\rho}^{\rho}\}$ linearly independent, then sF-G has a set of e.d. $\{(s-\alpha)^{d_1},\dots,(s-\alpha)^{d_\rho}\}$.

(ii) If there exist maximal chain of linearly independent vectors $\{\hat{\underline{x}}^i_1,\dots,\hat{\underline{x}}^i_{d'_i}\}$, $i\in\underset{\sim}{\rho}'$, such that

$$F\hat{\underline{x}}^i_j = \hat{\alpha}G\hat{\underline{x}}^i_j+G\hat{\underline{x}}^i_{j-1},\ j\in\underset{\sim}{d}'_i,\ \hat{\underline{x}}^i_o=0 \tag{4.19}$$

with the vector set $\{\hat{\underline{x}}^1_1,\dots,\hat{\underline{x}}^1_{d'_1};\dots;\hat{\underline{x}}^{\rho'}_1,\dots,\hat{\underline{x}}^{\rho'}_{d'_{\rho'}}\}$ linearly independent, then $F-\hat{s}G$ has a set of e.d. $\{(\hat{s}-\hat{\alpha})^{d'_1},\dots,(\hat{s}-\hat{\alpha})^{d'_{\rho'}}\}$. □

By Propositions (4.1) and (4.2) a characterization of the set of e.d. of a regular pencil may be derived. Thus, we have:

Theorem (4.1): (i) The regular pencil sF-G has an e.d. $(s-\alpha)^d$, if and only if there exists a maximal chain of linearly independent vectors $\{\underline{x}_1,\underline{x}_2,\dots,\underline{x}_d\}$ such that

$$G\underline{x}_i = \alpha F\underline{x}_i+F\underline{x}_{i-1},\ i\in\underset{\sim}{d},\ \underline{x}_o=\underline{0} \tag{4.20}$$

or equivalently

$$\begin{bmatrix} G-\alpha F & 0 & \dots & 0 & 0 \\ -F & G-\alpha F & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & -F & G-\alpha F \end{bmatrix}\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_i \end{bmatrix}=0,\ i\in\underset{\sim}{d} \tag{4.21}$$

(ii) The regular pencil $F-\hat{s}G$ has an e.d. $(\hat{s}-\hat{\alpha})^{d'}$, if and only if there exists a maximal chain of linearly independent vectors $\{\hat{\underline{x}}_1,\dots,\hat{\underline{x}}_d\}$ such that

$$F\hat{\underline{x}}_i = \hat{\alpha}G\hat{\underline{x}}_i+G\hat{\underline{x}}_{i-1},\ i\in\underset{\sim}{d}',\ \hat{\underline{x}}_o=\underline{0}$$

or equivalently

$$\begin{bmatrix} F-\hat{\alpha}G & 0 \; \dots & 0 & 0 \\ -G & F-\hat{\alpha}G \dots & 0 & 0 \\ \vdots & \ddots & & \vdots \\ 0 & 0 \; \dots & -G & F-\hat{\alpha}G \end{bmatrix} \begin{bmatrix} \hat{\underline{x}}_1 \\ \hat{\underline{x}}_2 \\ \vdots \\ \hat{\underline{x}}_i \end{bmatrix} = 0, \; i \in \underset{\sim}{d}' \qquad (4.22)$$

□

By noting that e.d. of the type s^p and $\hat{s}^q$ are special cases of $(s-\alpha)^p$ for $\alpha=0$ and $(\hat{s}-\hat{\alpha})^q$ for $\hat{\alpha}=0$ respectively and taking into account the duality of e.d. established in the previous section we have the following two results:

<u>Corollary (4.1)</u>: The regular pencil sF-G $(F-\hat{s}G)$ has a zero e.d. (infinite e.d.) s^p if and only if there exists a maximal linearly independent vector chain $\{\underline{x}_1,\dots,\underline{x}_p\} \in \mathbb{R}^n$ such that

$$G\underline{x}_i = F\underline{x}_{i-1}, \; i \in \underline{p}, \; \underline{x}_o = 0 \qquad (4.23)$$

or equivalently

$$\begin{bmatrix} G & 0 \; \dots & 0 & 0 \\ -F & G \; \dots & 0 & 0 \\ \vdots & \ddots & & \vdots \\ 0 & 0 \; \dots & G & 0 \\ 0 & 0 & -F & G \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_{i-1} \\ \underline{x}_i \end{bmatrix} = 0, \; i \in \underline{p} \qquad (4.24)$$

□

<u>Corollary (4.2)</u>: The regular pencil sF-G $(F-\hat{s}G)$ has an infinite e.d. (zero e.d.) $\hat{s}^q$ if and only if there exists a maximal linearly independent vector chain $\{\underline{x}_1,\underline{x}_2,\dots,\underline{x}_q\} \in \mathbb{R}^n$ such that

$$F\underline{x}_i = G\underline{x}_{i-1}, \; i \in \underline{q}, \; \underline{x}_o \; \underline{0} \qquad (4.25)$$

or equivalently

$$\begin{bmatrix} F & 0 & \dots & 0 & 0 \\ -G & F & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & F & 0 \\ 0 & 0 & \dots & -G & F \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_{i-1} \\ \underline{x}_i \end{bmatrix} = 0, \; i \in \underline{q} \qquad (4.26)$$

□

The above results suggest that the finite and infinite e.d. of a regular pencil may be defined independently of the Smith algorithm originally used for their definition. In the following, the problem of computing the degrees of the e.d. as well as of the vector chains associated with them is examined.

Let sF-G be an n×n regular pencil. For every $\alpha\in\mathbb{C}$, we may define the following matrices associated with the pair (F,G):

$$P_\alpha^1(F,G) = G-\alpha F \in \mathbb{C}^{n\times n}, \; P_\alpha^2(F,G) = \begin{bmatrix} G-\alpha F & 0 \\ -F & G-\alpha F \end{bmatrix} \in \mathbb{C}^{2n\times 2n}$$

$$\dots, P_\alpha^i(F,G) = \begin{bmatrix} G-\alpha F & 0 & \dots & 0 & 0 \\ -F & G-\alpha F & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & G-\alpha F & 0 \\ 0 & 0 & \dots & -F & G-\alpha F \end{bmatrix} \in \mathbb{C}^{in\times in} \quad i=1,2,\dots \qquad (4.27)$$

For the case of s=∞ we define the matrices $P_\infty^i(F,G)$ by

$$P_\infty^1(F,G) = F \in \mathbb{R}^{n\times n}, \; P_\infty^2(F,G) = \begin{bmatrix} F & 0 \\ -G & F \end{bmatrix} \in \mathbb{R}^{2n\times 2n}, \dots$$

$$\dots, P_\infty^i(F,G) = \begin{bmatrix} F & 0 & \dots & 0 & 0 \\ -G & F & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & F & 0 \\ 0 & 0 & \dots & -G & F \end{bmatrix} \in \mathbb{R}^{in\times in} \quad i=1,2,\dots \qquad (4.28)$$

The rank of the matrix $P_\alpha^i(F,G)$ $(P_\infty^i(F,G))$ shall be denoted by ρ_α^i (ρ_∞^i)

and the corresponding nullity (rank deficiency) shall be denoted by n_α^i (n_∞^i) respectively, note that $n_\alpha^i = ni - \rho_\alpha^i$. Matrices of the type $P_\alpha^i(F,G)$, $P_\infty^i(F,G)$ shall be referred to as <u>i-th order α-, ∞-Toeplitz matrices</u> of (F,G) respectively; some important notions associated with them are defined next.

<u>Definition (4.1)</u>: (i) We define as the <u>index of annihilation of (F,G) at $s=\alpha$</u> ($s=\infty$), the smallest integer τ_α (τ_∞) for which $n_\alpha^{\tau_\alpha} = n_\alpha^{\tau_\alpha+1}$ ($n_\infty^{\tau_\infty} = n_\infty^{\tau_\infty+1}$).
(ii) Let $\mathcal{N}_\alpha^i = \mathcal{N}_r(P_\alpha^i(F,G))$ (α is finite or infinite) and let $N_\alpha^i \in \mathbb{C}^{ni \times n_\alpha^i}$ be a basis matrix for $\mathcal{N}_\alpha^i$ (a right annihilator of $P_\alpha^i(F,G)$). Let $M_\alpha^i \in \mathbb{C}^{n \times n_\alpha^i}$ be the submatrix of N_α^i which is made up from the last n rows of N_α^i. The vector space $\mathcal{M}_\alpha^i = \text{col.span}\{M_\alpha^i\}$ is defined as the <u>i-th generalized null space</u> of (F,G) at $s=\alpha$.

The above notions are natural extensions of the notions of index of annihilation and generalized nullspace defined on linear operators, to the case of a pair of linear operators. The spectral analysis of regular matrix pencils heavily relies on those two notions. We start off by stating the following preliminary result.

<u>Proposition (4.4)</u>: Let $sF-G$ be a $n \times n$ regular pencil. The following properties hold true:

(i) The i-th generalized nullspaces of (F,G) at $s=\alpha$, $i=1,2,\ldots$ is non-trivial ($\neq\{0\}$) if and only if $G-\alpha F$ is singular.

(ii) The i-th generalized nullspaces of (F,G) at $s=\alpha$, $i=1,2,\ldots$ are nested, i.e. $\mathcal{M}_\alpha^i \subseteq \mathcal{M}_\alpha^{i+1}$ for $\forall$ $i=1,2,\ldots$ and the chain $\mathcal{M}_\alpha^1, \mathcal{M}_\alpha^2, \ldots, \mathcal{M}_\alpha^i$ has a maximal element $\mathcal{M}_\alpha^*$.

<u>Proof</u>

There are two possible cases: $G-\alpha F$ nonsingular and $G-\alpha F$ singular. If $|G-\alpha F| \neq 0$, then because of the upper triangular structure of $P_\alpha^i(F,G)$, $P_\alpha^i(F,G)$ are nonsingular and thus, $\mathcal{N}_\alpha^i = \{0\}$ and $\mathcal{M}_\alpha^i = \{0\}$. The upper triangular structure of $P_\alpha^i(F,G)$ and the fact that $G-\alpha F$ is on the diagonal guarantees that the $N_\alpha^i \neq \{0\}$. In order to show that $M_\alpha^i \neq \{0\}$ we apply induction. Thus, for i=1, $\mathrm{P}_\alpha^1(\mathrm{F},\mathrm{G}) = \mathrm{G}-\alpha\mathrm{F}$ and since $|\mathrm{G}-\alpha\mathrm{F}|=0$ there exists a right annihilator $\mathrm{N}_\alpha = \mathrm{X}_1^1 \in \mathbb{C}^{n\times\eta_\alpha^1}$ of $\mathrm{G}-\alpha\mathrm{F}$, i.e.

$$[\mathrm{G}-\alpha\mathrm{F}]\mathrm{X}_1^1 = 0,\ \mathrm{X}_1^1 \neq 0$$

Thus, $M_\alpha^1 \neq \{0\}$. For i=2, $\mathrm{P}_\alpha^2(\mathrm{F},\mathrm{G})$ becomes

$$\mathrm{P}_\alpha^2(\mathrm{F},\mathrm{G}) = \begin{bmatrix} \mathrm{G}-\alpha\mathrm{F} & 0 \\ -\mathrm{F} & \mathrm{G}-\alpha\mathrm{F} \end{bmatrix}$$

Clearly, the nullity of $\mathrm{P}_\alpha^2(\mathrm{F},\mathrm{G})$ is greater or equal to the nullity of $\mathrm{P}_\alpha^1(\mathrm{F},\mathrm{G})$ since the columns of the matrix $\bar{\mathrm{N}}_\alpha^2$, where

$$\bar{\mathrm{N}}_\alpha^2 = \begin{array}{cc} \leftarrow\eta_\alpha^1\rightarrow & \\ \left[\begin{array}{c} 0 \\ \hdashline \mathrm{X}_1^1 \end{array}\right] & \begin{array}{c} \updownarrow n \\ \updownarrow n \end{array} \end{array}$$

are in N_α^2. Starting from $\bar{\mathrm{N}}_\alpha^2$ we may complete the basis of N_α^2 and obtain a right annihilator of $\mathrm{P}_\alpha^2(\mathrm{F},\mathrm{G})$ of the form

$$\mathrm{N}_\alpha^2 = \left[\begin{array}{c:c} 0 & \mathrm{X}_1^1 \\ \hdashline \mathrm{X}_1^1 & \mathrm{X}_2^2 \end{array}\right] \in \mathbb{C}^{2n\times\eta_\alpha^2}$$

Since the columns of X_1^1 are linearly independent, $\mathrm{sp}\{[\mathrm{X}_1^1,\mathrm{X}_2^2]\} \neq \{0\}$ and thus $M_\alpha^2 \neq \{0\}$; furthermore, $\mathrm{sp}\{[\mathrm{X}_1^1]\} \subseteq \mathrm{sp}\{[\mathrm{X}_1^1,\mathrm{X}_2^2]\}$. Assume now that N_α^i is a right annihilator of $\mathrm{P}_\alpha^i(\mathrm{F},\mathrm{G})$, where

$$\mathrm{N}_\alpha^i = \left[\begin{array}{c:c:c:c:c} & & & 0 & \mathrm{X}_i^1 \\ \hdashline & & 0 & \iddots & \vdots \\ \hdashline & 0 & \mathrm{X}_3^1 & \cdots & \mathrm{X}_i^{i-2} \\ \hdashline 0 & \mathrm{X}_2^1 & \mathrm{X}_3^2 & \cdots & \mathrm{X}_i^{i-1} \\ \hdashline \mathrm{X}_1^1 & \mathrm{X}_2^2 & \mathrm{X}_3^3 & \cdots & \mathrm{X}_i^i \end{array}\right] \in \mathbb{C}^{ni\times\eta_\alpha^i}$$

and $M_\alpha^i = \mathrm{sp}\{[\mathrm{X}_1^1,\ldots,\mathrm{X}_i^i]\} \neq \{0\}$. Then we may write that

$$P_\alpha^{i+1}(F,G) = \left[\begin{array}{c|c} G-\alpha F & 0 \;\dots\; 0 \\ \hline -F & \\ 0 & \\ \vdots & P_\alpha^i(F,G) \\ 0 & \end{array}\right]$$

and clearly the columns of $\bar{N}_\alpha^i$, where

$$\bar{N}_\alpha^i = \begin{array}{cc} \left[\begin{array}{c} 0 \\ \hline N_\alpha^i \end{array}\right] & \begin{array}{c} \updownarrow n \\ \updownarrow ni \end{array} \\ \leftarrow n_\alpha^i \rightarrow & \end{array}$$

are independent and in N_α^{i+1}. If $n_\alpha^{i+1} > n_\alpha^i$, then the columns of $\bar{N}_\alpha^i$ may be completed to a basis of N_α^{i+1}. A right annihilator of $P_\alpha^{i+1}(F,G)$ is then of the form

$$N_\alpha^{i+1} = \left[\begin{array}{c|c} 0 & X_{i+1}^1 \\ \hline N_\alpha^i & \vdots \\ & X_{i+1}^{i+1} \end{array}\right]$$

Thus, $M_\alpha^{i+1} = \text{sp}\{[X_1^1, \dots, X_i^i, X_{i+1}^{i+1}]\} \neq \{0\}$ and $M_\alpha^i \subseteq M_\alpha^{i+1}$. If $n_\alpha^{i+1} = n_\alpha^i$, then $\bar{N}_\alpha^i$ is a right annihilator of $P_\alpha^{i+1}(F,G)$ and $M_\alpha^i = M_\alpha^{i+1}$.

Note that $M_\alpha^i \in \mathbb{C}^n$ and thus there is an upper bound for dim M_α^i; the nesting property $M_\alpha^i \subseteq M_\alpha^{i+1}$ implies then the existence of a maximal element M_α^*. □

The following Corollary is readily established from the above result.

Corollary (4.3): Let sF-G be an n×n regular pencil and let G-αF be singular for $\alpha \in \mathbb{C}$. There exists a basis for N_α^i of the type

$$N_\alpha^i = \left[\begin{array}{cccccc|c} 0 & 0 & 0 & \dots & 0 & & X_i^1 \\ \hline \vdots & & & & \cdot & & X_i^2 \\ \hline & & & \cdot & & & \vdots \\ & & 0 & & & & \\ \hline & 0 & X_3^1 & & \dots & & X_i^{i-2} \\ \hline 0 & X_2^1 & X_3^2 & & \dots & & X_i^{i-1} \\ \hline X_1^1 & X_2^2 & X_3^3 & & \dots & & X_i^i \end{array}\right] = \begin{array}{cc} \left[\begin{array}{c|c} 0 \;\dots\; 0 & X_i^1 \\ \hline & \vdots \\ N_\alpha^{i-1} & X_i^{i-1} \\ & X_i^i \end{array}\right] & \updownarrow n \end{array} \tag{4.29}$$

where N_{α}^{i-1} is a basis matrix for $\mathcal{N}_{\alpha}^{i-1}$, X_1^1 a basis matrix for $\mathcal{N}_{\alpha}^{1}$ and the columns of X_j^1, j=2,3,...i are in $\mathcal{N}_{\alpha}^{1}$.

□

Remark (4.1): The subspaces $\mathcal{N}_{\alpha}^{i}$, i=1,2,... are nontrivial ($\neq\{0\}$), if and only if $G-\alpha F$ is singular; the matrix $G-\alpha F$, however, is nonsingular if and only if the regular pencil sF-G has a set of e.d. at $s=\alpha$. Thus, it follows that the subspaces $\mathcal{N}_{\alpha}^{i}$ ($\mathcal{N}_{\infty}^{i}$) are nontrivial if and only if sF-G has a set of e.d. at $s=\alpha$ ($s=\infty$).

4.3 Segre, Weyr characteristics, and the piecewise arithmetic progression sequences of (F,G) at $s=\alpha$ (∞)

In this section, attention is focussed on the investigation of the links between the properties of $\mathcal{N}_{\alpha}^{i}$ ($\mathcal{N}_{\infty}^{i}$) and the structure of the e.d. set of sF-G at $s=\alpha$. The result of the present analysis is the introduction of two alternative procedures for the computation of the set of degrees of the e.d. at $s=\alpha$ ($s=\infty$) of $sF-\hat{s}G$, rather than the Smith form based definition of them. The results will be presented for $s=\alpha\in\mathbb{C}$, whereas the case of $s=\infty$ is similar (simply use $P_{\infty}^{i}(F,G)$ instead of $P_{\alpha}^{i}(F,G)$). We start off by giving the following useful result.

Lemma (4.1): Let $H=\text{diag}\{H_1,\dots,H_{\nu}\}\in\mathbb{R}^{n\times n}$, where

$$H_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \vdots & & & & \ddots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \in\mathbb{R}^{d_i\times d_i}$$

$d_1\le d_2\le\dots\le d_{\nu}$, $\mathcal{N}_k=\mathcal{N}_r\{H^k\}$ and $n_k=\dim\mathcal{N}_k$. Then, the following properties hold true:

(i) The set of subspaces $\mathcal{N}_i$ is nested, i.e. $\mathcal{N}_i\subseteq\mathcal{N}_{i+1}$ and the chain has a maximal element $\mathcal{N}^*=\mathbb{R}^d$, where $d=\sum_{i=1}^{n} d_i$.

(ii) The smallest index τ for which $\mathcal{N}_{\tau-1}\subset\mathcal{N}_{\tau}=\mathcal{N}^*$ is called the index of annihilation of H and $\tau=\max\{d_i, i\in\underset{\sim}{\nu}\}$.

(iii) $\eta_k = \sum_{i=1}^{\underline{n}} \eta_i^k$ where $\eta_i^k = d_i$ if $k \geq d_i$ and $\eta_i^k = k$ if $k < d_i$.

<u>Proof</u>

It is clear that $H^k = \mathrm{diag}\{H_1^k, \ldots, H_\nu^k\}$ for all $k = 1, 2, \ldots$ Because of the block-diagonal structure of H^k, $N_r\{H^k\}$ is defined as a direct sum of the null spaces corresponding to the H_i^k matrices. Note, that if $k \geq d_i$, then $H_i^k = 0$ and if $k < d_i$, then

$$H_i^k = \left[\begin{matrix} 0 & \cdots 0 & 1 & 0 \cdots & \cdots 0 \\ 0 & \cdots 0 & 0 & 1 & \vdots \\ & & & \ddots & \\ 0 & \cdots 0 & 0 & \cdots & 1 \\ \vdots & \vdots & & & \vdots \\ 0 \ldots & 0 & 0 & \ldots & 0 \end{matrix}\right] \begin{matrix} \big\uparrow \\ d_i - k \\ \big\downarrow \\ \\ \\ \end{matrix} \tag{4.30a}$$

and thus a basis matrix for $N_r\{H_i^k\}$ is defined by

$$N_{d_i}^k = I_{d_i} \text{ if } k \geq d_i \text{ and } N_{d_i}^k = \begin{bmatrix} I_k \\ \hline 0 \end{bmatrix} \text{ if } k < d_i \tag{4.30b}$$

By inspection of the above it is clear that $N_n\{H_i^k\} \subseteq N_r\{H_i^{k+1}\}$ with equality holding only when $k \geq d_i$. Since $N_r\{H^k\}$ may be expressed as a direct sum of the nullspaces associated with the H_i^k blocks, it follows that $N_i \subseteq N_{i+1}$. Clearly the smallest index for which $N_\tau = N_{\tau+1}$ is $\tau = \max\{d_i, i \in \underset{\sim}{\nu}\}$, because this is minimal value of j for which $H^j = 0$ and thus $N_\tau = \mathbb{R}^d$, where $d = \sum_{i=1}^{\nu} d_i$. Because $\eta_k = \dim N_r\{H^k\} = \sum_{i=1}^{\nu} \dim N_r\{H_i^k\}$ and the special structure of H_i, we have that if $k < d_i$, then $\eta_i^k = \dim N_r\{H_i^k\} = k$ and if $k \geq d_i$, then $\eta_i^k = d_i$ since $H_i^k = 0$ and $\dim N_r\{H_i^k\} = d_i$.

□

<u>Theorem (4.3)</u>: Let sF-G be an n×n regular pencil, let G-αF be singular at s=α and let $\{(s-\alpha)^{d_1}, \ldots, (s-\alpha)^{d_\nu}\}$ be the set of e.d. at s=α. The following properties hold true:

(i) $\eta_\alpha^k = \dim N_\alpha^k = \sum_{i=1}^{\nu} \eta_i^k$, where $\eta_i^k = d_i$, if $k \geq d_i$ and $\eta_i^k = k$, if $k < d_i$.

(ii) $\dim M_\alpha^k = \dim N_\alpha^k = \eta_\alpha^k$.

(iii) The smallest integer τ for which $M_\alpha^{\tau-1} \subset M_\alpha^\tau = M_\alpha^{\tau+1} = M_\alpha^*$, is the index of annihilation τ_α of (F,G) at s=α; $\tau_\alpha = \max\{d_i, i \in \underset{\sim}{\nu}\}$ and $\dim M_\alpha^* = d = \sum_{i=1}^{\nu} d_i$.

Proof

(i) It is clear that if $R,Q \in \mathbb{C}^{n\times n}$, $|R|,|Q| \neq 0$, then $\mathrm{rank}\{P^{i}(F,G)\} = $ $=\mathrm{rank}\{\mathrm{diag}\{R,\dots,R\}P_{\alpha}^{i}(F,G)\mathrm{diag}\{Q,\dots,Q\}\}$ and thus, the rank property may be tested by using the Weierstrass canonical pair (F_w,G_w). Assume $G_w-\alpha F_w$ in the Weierstrass form, defined by (4.11); the blocks $J_{\tau_i}(\beta_i)-\alpha I_{\tau_i}$, $I_{q_i}-\alpha H_{q_i}$ are non-singular, since $\beta_i \neq \alpha$, and thus we may write

$$G_w-\alpha F_w = \begin{bmatrix} T & 0 \\ 0 & H \end{bmatrix}$$

where T is the block diagonal matrix made up from the non-singular blocks and $H=\mathrm{diag}\{H_1,\dots,H_\nu\}$, where $H_i=J_{d_i}(\alpha)-\alpha I_{d_i} \in \mathbb{R}^{d_i\times d_i}$. Note that $\eta_{\alpha}^{k}=\dim\mathcal{N}_r\{P^{k}(F_w,G_w)\}$ and thus η_{α}^{k} is defined by finding the maximal number of linearly independent solutions of the equation

$$P_{\alpha}^{k}(F_w,G_w)\underline{\tilde{x}} = \underline{0}$$

Let $\sum_{i=1}^{\nu} d_i = d$ and $\sigma=n-d$. By partitioning $\underline{\tilde{x}}$ according to the partitioning of $P_{\alpha}^{k}(F_w,G_w)$ we have that the following conditions must be satisfied for every k

$$\begin{bmatrix} T & 0 & & & & & & \\ 0 & H & & & \bigcirc & & & \\ D & 0 & T & 0 & & & & \\ 0 & I_d & 0 & H & & & & \\ & & & & \ddots & & & \\ & \bigcirc & & & D & 0 & T & 0 \\ & & & & 0 & I_d & 0 & H \end{bmatrix} \begin{bmatrix} \underline{\tilde{x}}_1 \\ \underline{\hat{x}}_1 \\ \underline{\tilde{x}}_2 \\ \underline{\hat{x}}_2 \\ \vdots \\ \underline{\tilde{x}}_k \\ \underline{\hat{x}}_k \end{bmatrix} = 0, \; D \in \mathbb{R}^{\sigma\times\sigma} \qquad (4.31)$$

It may be readily verified that the above equation is equivalent to the following two equations

$$\left[\begin{array}{c:c:c:c} T & & & \\ \hdashline D & T & & \bigcirc \\ \hdashline & & \ddots & \\ \hdashline \bigcirc & & D & T \end{array}\right]\left[\begin{array}{c} \tilde{\underline{x}}_1 \\ \hdashline \tilde{\underline{x}}_2 \\ \hdashline \vdots \\ \hdashline \tilde{\underline{x}}_k \end{array}\right] = 0 \qquad (4.32a)$$

and

$$\left[\begin{array}{c:c:c:c} H & & & \\ \hdashline I_d & H & & \bigcirc \\ \hdashline & & \ddots & \\ \hdashline \bigcirc & & I_d & H \end{array}\right]\left[\begin{array}{c} \hat{\underline{x}}_1 \\ \hdashline \hat{\underline{x}}_2 \\ \hdashline \vdots \\ \hdashline \hat{\underline{x}}_k \end{array}\right] = 0 \qquad (4.32b)$$

Given that T is nonsingular, eqn(4.32a) has as the only solution the $\tilde{\underline{x}}_1=\tilde{\underline{x}}_2=\ldots=\tilde{\underline{x}}_k=\underline{0}$; thus, η_α^k is defined by the number of independent solutions of (4.32b). Eqn(4.32b) is equivalent to

$$H\hat{\underline{x}}_1 = 0,\ \hat{\underline{x}}_1 = H\hat{\underline{x}}_2, \ldots, \hat{\underline{x}}_{k-1} = H\hat{\underline{x}}_k \qquad (4.32c)$$

or equivalently

$$\hat{\underline{x}}_1 = H\hat{\underline{x}}_2,\ \hat{\underline{x}}_2 = H\hat{\underline{x}}_3, \ldots, \hat{\underline{x}}_{k-1} = H\hat{\underline{x}}_k,\ H^k\hat{\underline{x}}_k = 0 \qquad (4.32d)$$

Let $\hat{X}_k$ be a right annihilator of H^k. Conditions (4.32d) suggest that matrices $\hat{X}_i$ may be defined by

$$\hat{X}'_{k-1} = H\hat{X}_k,\ \hat{X}'_{k-2} = H^2\hat{X}_k, \ldots,\ \hat{X}'_{k-j} = H^j\hat{X}_k \qquad (4.32e)$$

and thus

$$\hat{X}'_{k-j} = \left[\begin{array}{c} \hat{X}_k^{k-j]} \\ \hdashline 0 \end{array}\right] \begin{array}{c} \\ \updownarrow j \end{array} \qquad (4.43f)$$

where $\hat{X}_k^{k-j]}$ is the submatrix of $\hat{X}_k$ formed from the last k-j rows. Clearly the row space of the matrices

$$\hat{X} = \begin{bmatrix} \hat{X}_1' \\ \vdots \\ \hat{X}_{k-1}' \\ \hat{X}_k \end{bmatrix} \quad , \quad X = \left[\begin{array}{c} 0 \\ \hdashline \hat{X}_1' \\ \hdashline \vdots \\ \hdashline 0 \\ \hdashline \hat{X}_{k-1}' \\ \hdashline 0 \\ \hdashline \hat{X}_k \end{array}\right] \begin{array}{l} \updownarrow \sigma \\ \updownarrow d \\ \\ \updownarrow \sigma \\ \updownarrow d \\ \updownarrow \sigma \\ \updownarrow d \end{array} \tag{4.32g}$$

are equal to the row space of $\hat{X}_k$. X, however, is a right annihilator of $P_\alpha^k(F_w,G_w)$ and thus $\eta_\alpha^k=\dim\mathcal{N}_\alpha^k$ is equal to the number of independent rows, or columns of $\hat{X}_k$, or otherwise equal to $\dim\mathcal{N}_r\{H^k\}$. Thus, $\dim\mathcal{M}_\alpha^k=\eta_\alpha^k=$ $=\dim\mathcal{N}_r\{H^k\}$ and by Lemma (4.1) the proof of part (i) and part (ii) is established.

(iii) Since X, as defined by (4.32g), is a right annihilator of $P_\alpha^k(F_w,G_w)$ and $\hat{X}_k$ is a right annihilator of H^k (full rank matrix), $\mathcal{N}_\alpha^i$ and thus $\mathcal{M}_\alpha^i$ become maximal when $\mathcal{N}_r\{H^k\}$becomes maximal. By Lemma (4.1) $\mathcal{N}_r\{H^k\}$ becomes maximal for $k\geq\max\{d_i, i\in\underset{\sim}{\nu}\}$; the smallest integer τ_α for which $\mathcal{N}_\alpha^{\tau_\alpha-1}\subset\mathcal{N}_\alpha^{\tau_\alpha}=\mathcal{N}_\alpha^{\tau_\alpha+1}$ is clearly $\tau_\alpha=\max\{d_i, i\in\underset{\sim}{\nu}\}$ for which $\mathcal{N}_r\{H^{\tau_\alpha-1}\}\subset\mathcal{N}_r\{H^{\tau_\alpha}\}=$ $=\mathcal{N}_r\{H^{\tau_\alpha+1}\}$. For $k\geq\max\{d_i, i\in\underset{\sim}{\nu}\}$, $H^k=0$ and thus $\hat{X}_k=I_d$, $d=\sum_{i=1}^{\nu} d_i$. Thus, $\dim\mathcal{M}_\alpha^*=d=\sum_{i=1}^{\nu} d_i$. □

The above result establishes the links between the spaces $\mathcal{N}_\alpha^i$, $\mathcal{M}_\alpha^i$ and the degrees of the e.d. at s=α and thus provides the basis for the definition of the degrees of the e.d. at s=α without resorting to the construction of the Smith form. The maximal subspace $\mathcal{M}_\alpha^*$ plays an important role in the determination of the chains of vectors associated with the set of e.d. and shall be referred to as the <u>maximal generalized nullspace</u> of (F,G) at s=α. Before we examine the geometry of the e.d. set of (F,G) at s=α, a number of important results characterising the e.d. as numerical invariants of the pair (F,G) are given first. Some useful notation is introduced first. Let σ_i be the multiplicity of the e,d, $(s-\alpha)^{d_i}$; the ordered pair (d_i,σ_i) characterises such e.d. The ordered set $I=\{(d_1,\sigma_1),\ldots,(d_\rho,\sigma_\rho),d_1<\ldots<d_\rho\}$ characterises the totality of the e.d. of sF-G at s=α and shall be referred to as the <u>index set</u> of (F,G) at s=α.

<u>Corollary (4.4)</u>: Let sF-G be a regular pencil, G-αF singular and let $\{(d_1,\sigma_1),\ldots,(d_\rho,\sigma_\rho);d_1<\ldots<d_\rho\}$ be the index set of (F,G) at s=α. The dimensions η_k of the subspaces N_α^k of (F,G) at s=α satisfy the following properties:

(i) If $d_i\le k<d_{i+1}$, then

$$\eta_k = \sum_{j=1}^{i}\sigma_j d_j+k(\sigma_{i+1}+\ldots+\sigma_\rho) \tag{4.33}$$

and $\eta_k=k(\sigma_1+\ldots+\sigma_\rho)$ if $k<d_1$ and $\eta_k=\sum_{i=1}^{\rho}\sigma_i d_i$ if $k\ge d_\rho$.

(ii) For all k=0,1,2,..., then $\eta_{k+1}-\eta_k=\sigma_i+\ldots+\sigma_\rho$, where $\sigma_i,\ldots,\sigma_\rho$ are the multiplicities of the e.d. $(s-\alpha)^{d_i}$ for which $k<d_i<\ldots<d_\rho$ and $\eta_o=0$. For $k\ge d_\rho$, $\eta_{k+1}-\eta_k=0$.

<u>Proof</u>

(i) The proof of part (i) follows immediately from part (i) of Theorem (4.3) by adopting the notation introduced before.

(ii) If $d_1<\ldots<d_i<\ldots<d_\rho$ are the degrees of the e.d. at s=α then the following possibilities exist

(a) $d_{i-1}\le k<k+1<d_i$. Then, by (4.33) we have

$$\eta_{k+1}-\eta_k = \sum_{j=1}^{i-1}\sigma_j d_j+(k+1)(\sigma_1+\ldots+\sigma_\rho) - \sum_{j=1}^{i-1}\sigma_j d_j-k(\sigma_1+\ldots+\sigma_\rho)$$

$$=\sigma_i+\ldots+\sigma_\rho$$

(b) $d_{i-1}\le k<k+1=d_i$. Then, by (4.33) we have

$$\eta_{k+1}-\eta_k = \sum_{j=1}^{i}\sigma_j d_j+d_i(\sigma_{i+1}+\ldots+\sigma_\rho) - \sum_{j=1}^{i-1}\sigma_j d_j-(d_i-1)(\sigma_i+\ldots+\sigma_\rho)$$

$$=\sigma_i+\ldots+\sigma_\rho$$

(c) $d_\rho\le k<k+1$. Then, by part (i) we have $\eta_{k+1}=\eta_k$ and $\eta_{k+1}-\eta_k=0$.

(d) $k=0$. Then, $\eta_o=0$ and $\eta_1=\sigma_1+\ldots+\sigma_\rho$. Thus, $\eta_1-\eta_o=\sigma_1+\ldots+\sigma_\rho$. □

The above corollary provides the means for the computation of the e.d. structure of (F,G) at $s=\alpha$ by using the dimensions of the subspaces $\mathcal{N}_\alpha^k$, or the nullities of the $P_\alpha^k(F,G)$ matrices.

Remark (4.2): The differences $\eta_{k+1}-\eta_k$ provide the following information about the e.d. structure of sF-G at $s=\alpha$.

(i) $\eta_1-\eta_o=\sigma_1+\ldots+\sigma_\rho$ is the number of e.d. at $s=\alpha$.

(ii) The smallest index k for which $\eta_{k+1}-\eta_k=0$ gives the index of annihilation τ_α, which is also equal to the maximal degree d_ρ.

(iii) The difference $\eta_{k+1}-\eta_k$ defines the number of e.d. with degrees higher than k.

The computation of the complete set of degrees and multiplicities makes use of the following result.

Corollary (4.5): For all $k=1,2,\ldots$, the numbers η_k satisfy the relationship

$$\eta_k \geq \frac{\eta_{k-1}+\eta_{k+1}}{2} \tag{4.34}$$

In particular, we have that

(i) Strict inequality holds if and only if k is the degree of an e.d. of (F,G) at $s=\alpha$.

(ii) Equality holds if and only if k is not the degree of an e.d.

Proof

Necessary and sufficient condition for the η_k to be nonzero is that sF-G has a set of e.d. at $s=\alpha$. Assume, then that $\{(d_1,\sigma_1),\ldots,(d_\rho,\sigma_\rho)\}$ be the set of ordered pairs of degrees and multiplicities. We may distinguish the following cases:

(a) $k>d_\rho$. Then, by Corollary (4.4) we have $\eta_{k+1}-\eta_k=0=\eta_k-\eta_{k-1}$ and thus (4.34) holds with the equality sign.

(b) $d_{i-1}\leq k-1<k<k+1\leq d_i$. Then, by Corollary (4.4) we have $\eta_k-\eta_{k-1}=\sigma_i+\ldots+\sigma_\rho$

and $\eta_{k+1}-\eta_k=\sigma_i+\ldots+\sigma_\rho$ from which the equality sign holds true.

(c) $d_{i-1}\le k-1<k=d_i<k+1\le d_{i+1}$. Then, by Corollary (4.4) we have $\eta_k-\eta_{k-1}=\sigma_i+\ldots+\sigma_\rho$, $\eta_{k+1}-\eta_k=\sigma_{i+1}+\ldots+\sigma_\rho$ and thus $\eta_k-\eta_{k-1}>\eta_{k+1}-\eta_k$ and the inequality sign holds true.

(d) If $k=1=d_1<2\le d_2$, then $\eta_1-\eta_o=\sigma_1+\ldots+\sigma_\rho$ and $\eta_2-\eta_1=\sigma_2+\ldots+\sigma_\rho$ and thus inequality holds true.

(e) If $k=1<2\le d_1$, then $\eta_1-\eta_o=\sigma_1+\ldots+\sigma_\rho$ and $\eta_2-\eta_1=\sigma_1+\ldots+\sigma_\rho$, from which the equality sign holds true.

The above analysis proves the necessity of part (i) and (ii). The sufficiency is proved by contradiction. Thus, let us first assume that inequality holds true for some k and that k is not the degree of an e.d. Then cases (a), (b), or (e) are the only possibilities for k. Clearly, in either of these cases equality holds and this leads to a contradiction. Similarly, if we assume that equality holds for some k which is the degree of an e.d., then the only possible cases are (c) and (d) which clearly imply that inequality holds true, thus contradicting our assumption. □

The above result results in a very important property of the non-decreasing sequence of natural numbers $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ For all k integers which are not degrees of e.d., the numbers η_k satisfy the <u>arithmetic progression relationship</u> (APR) $\eta_k=1/2(\eta_{k-1}+\eta_{k+1})$. For those values of k which coincide with degrees of e.d., the arithmetic progression relationship is violated since $\eta_k>1/2(\eta_{k-1}+\eta_{k+1})$. The sequence $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ is partitioned by those values of k which correspond to the degrees of the e.d. If $d_1<d_2<\ldots<d_\rho$ are the possible degrees of elementary divisors, then the numbers $\eta_{d_i},\eta_{d_i+1},\ldots,\eta_{d_{i+1}-1},\eta_{d_{i+1}}$ are elements of an <u>arithmetic progression sequence</u> (APS), since $\eta_k=1/2(\eta_{k-1}+\eta_{k+1})$ for $k=d_i+1,\ldots,d_{i+1}-1$; this relationship that holds in the $(d_i+1,\ldots,d_{i+1}-1)$ range of k values cannot be continued in the

$(d_{i-1}+1,\ldots,d_i-1)$, or $(d_{i+1}+1,\ldots,d_{i+2}-1)$ since $\eta_{d_i}>1/2(\eta_{d_i+1}+\eta_{d_i-1})$ and $\eta_{d_{i+1}}>1/2(\eta_{d_{i+1}+1}+\eta_{d_{i+1}-1})$. The number $\delta_{d_i}=(\eta_{d_i}-\eta_{d_i-1})-(\eta_{d_i+1}-\eta_{d_i})=$ $=2\eta_{d_i}-\eta_{d_i-1}-\eta_{d_i+1}$ is a measure of discontinuity, or deviation between the APR holding in $(d_{i-1}+1,\ldots,d_i-1)$ and $(d_i+1,\ldots,d_{i+1}-1)$ respectively. The sequence $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ therefore satisfies the arithmetic progression type relationships in finite sets of successive natural numbers; the only points where such relationships do not hold (but become inequalities) are the degrees of the e.d. Such a sequence will be referred to as <u>piecewise arithmetic progression sequence</u> (PAPS) of (F,G) at $s=\alpha$; the value of $k=d$, where there is a discontinuity in the APR, will be referred to as a <u>singular point</u> and the number δ_d will be called the <u>gap</u> of the sequence at $k=d$.

With these observations in mind we can state the following result relating the PAPS $\eta_o,\eta_1,..,\eta_i$ of (F,G) at $s=\alpha$ with the index set of (F,G) at $s=\alpha$.

<u>Proposition (4.5)</u>: Let $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ be the PAPS of (F,G) at $s=\alpha$. Then

(i) An index $k=d_i$ is a singular point of the sequence, if and only if d_i is the degree of an e.d. of sF-G at $s=\alpha$.

(ii) If $k=d_i$ is a singular point, then the gap δ_{d_i} at $k=d$ is equal to the multiplicity σ_{d_i} of the e.d. at $s=\alpha$ with degree d_i.

<u>Proof</u>

Part (i) follows immediately by Corollary (4.5). Since $k=d_i$ is the degree of an e.d., then by Corollary (4.4) we have that $\eta_k-\eta_{k-1}=$ $=\sigma_i+\sigma_{i+1}+\ldots+\sigma_\rho$ and $\eta_{k+1}-\eta_k=\sigma_{i+1}+\ldots+\sigma_\rho$ and thus $\delta_{d_i}=(\eta_k-\eta_{k-1})-(\eta_{k+1}-\eta_k)=\sigma_i$. □

By finding the singular points of the PAPS, $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ and the corresponding gaps, the index set of (F,G) at $s=\alpha$ is defined. The analysis presented so far leads to the following procedure for the determination of the degrees of the e.d. at s=α.

Piecewise arithmetic progression sequence diagram (PAPSD): Compute the numbers $\eta_o, \eta_1, \eta_2, \ldots, \eta_k, \ldots$, with $\eta_o=0$ until we find the first index τ for which $\eta_\tau = \eta_{\tau+1}$. Then, τ is the index of annihilation of (F,G) at s=α. Compute then the gaps of the PAPS, i.e.

$$\delta_i = 2\eta_i - \eta_{i-1} - \eta_{i+1} \quad \text{for } i=1,2,\ldots,\tau,\tau+1, \eta_o=0$$

and form a table of the following type: For every index i there is a value $\delta_i \geq 0$; if $\delta_i = 0$ a dot is placed below δ_i and if $\delta_i > 0$, then we create a column with asterisks below δ_i, with the number of asterisks being equal to the value of δ_i. This procedure is illustrated by the following diagram:

$$\begin{array}{lccccccccc}
\text{index:} & 1, & 2, & \ldots, & i-1, & i, & i+1, & \ldots, & \tau, & \tau+1 \\
\text{gap :} & \delta_1, & \delta_2, & \ldots, & \delta_{i-1}, & \delta_i, & \delta_{i+1}, & \ldots, & \delta_\tau, & \delta_{\tau+1} \\
 & \cdot & * & & \cdot & * & \cdot & & * & \cdot \\
 & & \vdots & & & \vdots & & & \vdots & \\
 & & * & & & \vdots & & & * & \\
 & & & & & * & & & &
\end{array}$$

Fig.(4.1)

The indices characterised by dots do not correspond to degrees of e.d., whereas those characterised by asterisks define the degrees of the e.d. The number of asterisks in a column gives the multiplicity of the e.d. whose degree is the corresponding index. Thus for instance, in the above diagram we have e.d. with degrees 2,i,τ; the multiplicities indicated by this diagram are 2,4,3 respectively.

The above procedure will be illustrated later on by an example. An alternative technique, which may be used for determining the index set of (F,G) at s=α is discussed next. The following procedure is a generalisation of the standard procedure used for finding the Jordan blocks of a square matrix A, which is based on the notions of Weyr and Segre characteristics and used a Ferrer's type diagram for the partitioning of natural numbers [Turn & Ait].

The extensions of the standard notions used in the analysis of the sI-A pencil to the case of sF-G regular pencil is considered first.

Definition (4.2): Let sF-G be a regular pencil, G-αF singular and let $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ be the nullities of $P^i_{\alpha}(F,G)$, $i=0,1,\ldots,i,\ldots$, where $P^o_{\alpha}(F,G)\overset{\Delta}{=}I_n$. We may define the following:

(i) The Segré characteristic of (F,G) with respect to α is defined as the set of the first nonzero differences of the powers to which the scalar factor (s-α) occurs in |sF-G| and in the H.C.F.s of its minors of descending order. Clearly, if $\{(d_1,\sigma_1),\ldots,(d_{\rho},\sigma_{\rho});\ d_1<\ldots<d_{\rho}\}$ is the index set at s=α of (F,G), then the Segre characteristic at s=α is defined by

$$S_{\alpha}(F,G) = \{\underset{\longleftarrow\sigma_{\rho}\longrightarrow}{d_{\rho},\ldots,d_{\rho}};\ldots;\underset{\longleftarrow\sigma_{i}\longrightarrow}{d_{i},\ldots,d_{i}};\ldots;\underset{\longleftarrow\sigma_{1}\longrightarrow}{d_{1},\ldots,d_{1}}\} \tag{4.35}$$

(ii) The set of the first nonzero successive differences in $\eta_o,\eta_1,\ldots,\eta_i,\ldots$ is defined as the Weyr characteristic of (F,G) at s=α and it is denoted by $\mathcal{W}_{\alpha}(F,G)$. Clearly, if τ is the index of annihilation of (F,G) at s=α, then

$$\mathcal{W}_{\alpha}(F,G) = \{\gamma_1=\eta_1-\eta_o,\gamma_2=\eta_2-\eta_1,\ldots,\gamma_{\tau}=\eta_{\tau}-\eta_{\tau-1}\} \tag{4.36}$$

The Weyr characteristic contains all the information we need to define the Segre characteristic, as it is shown by the following result.

Proposition (4.6): Let $S_{\alpha}(F,G)$,$\mathcal{W}_{\alpha}(F,G)$ be the Segré,Weyr characteristics respectively of sF-G at s=α, as denoted by (4.35) and (4.36), correspondingly. Then,

(i) $\gamma_j \geq \gamma_{j+1}$ for all $j=1,2,\ldots,\tau$ and $\gamma_{k+1}=0$ for all $k=\tau,\tau+1,\ldots$

(ii) The strict inequality $\gamma_k > \gamma_{k+1}$ holds true if and only if $k=d_i$, where d_i is a Segre index (element of $S_\alpha(F,G)$). The multiplicity σ_i of d_i is then defined by $\sigma_i=\gamma_{d_i}-\gamma_{d_i+1}$. □

The above result is an alternative presentation of Proposition (4.6) and thus its proof is omitted. The proposition suggests a method for computing $S_\alpha(F,G)$ from $W_\alpha(F,G)$, which is known as a Ferrer's diagram [Turn & Ait].

Ferrer's diagram: Let $W_\alpha(F,G)=\{\gamma_1,\gamma_2,\ldots,\gamma_{\tau_\alpha}\}$ be the Weyr characteristic of (F,G) at α. For every number γ_i create a row with asterisks. In each row we put as many asterisks as the numbers $\gamma_1,\gamma_2,\ldots,\gamma_{\tau_\alpha}$ respectively. Then we count the number of asterisks in each column. This gives us the elements of the Segré characteristic $S_\alpha(F,G)$ of (F,G) at s=α. An illustration of this diagram is shown below:

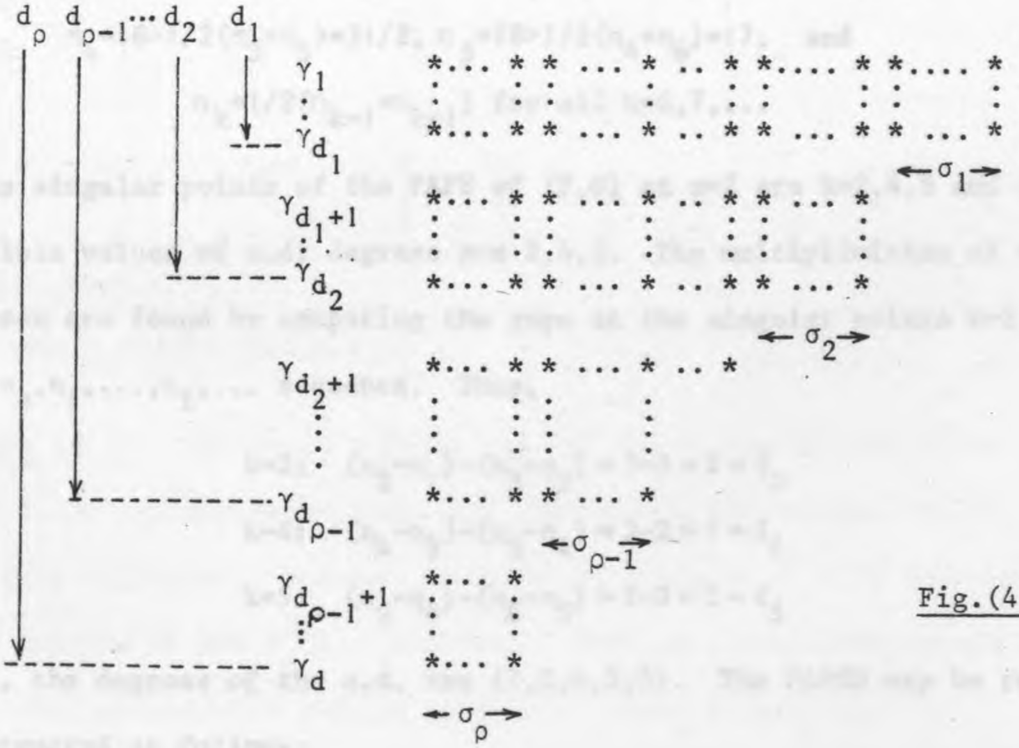

Fig.(4.2)

An example demonstrating the determination of the index set of the pair (F,G) at s=α, using both methods discussed before is given next.

Example (4.1): Let sF-G be a 20×20 regular pencil and let G-2F be rank deficient. The pencil sF-G has e.d. of the type $(s-2)^{d_i}$; the degrees and numbers of such e.d. may be found as follows: Let the ranks of $P_2^i(F,G)$ and the corresponding nullities be

$$\rho_o = 20,\ \rho_1 = 15,\ \rho_2 = 30,\ \rho_3 = 47,\ \rho_4 = 64,\ \rho_5 = 82,\ \rho_6 = 102$$

$$\eta_o = 0\ ,\ \eta_1 = 5\ ,\ \eta_2 = 10,\ \eta_3 = 13,\ \eta_4 = 16,\ \eta_5 = 18,\ \eta_6 = 18$$

Since $\eta_1=5$ we have 5 e.d. and let $d_1 \le d_2 \le d_3 \le d_4 \le d_5$ be their corresponding degrees. The index of annihilation is $\tau=5$ and thus $d_5=5$; because $\eta_5=18$, we also have $d_1+d_2+d_3+d_4+d_5=18$. The prediction of the distinct degrees may be achieved by using the piecewise arithmetic progression property of the sequence $\eta_0, \eta_1, \dots$, or by using the Ferrer diagram.

Piecewise arithmetic progression sequence analysis

$$\eta_1=5=1/2(\eta_o+\eta_2),\ \eta_2=10>1/2(\eta_1+\eta_3)=9,\ \eta_3=13=1/2(\eta_2+\eta_4)$$
$$\eta_4=16>1/2(\eta_3+\eta_5)=31/2,\ \eta_5=18>1/2(\eta_4+\eta_6)=17,\ \text{and}$$
$$\eta_k=1/2(\eta_{k-1}+\eta_{k+1}) \text{ for all } k=6,7,\dots$$

The singular points of the PAPS of (F,G) at s=2 are k=2,4,5 and thus the possible values of e.d. degrees are 2,4,5. The multiplicities of the degrees are found by computing the gaps at the singular points k=2,4,5 of the $\eta_o, \eta_1, \dots, \eta_i, \dots$ sequence. Thus,

$$k=2:\quad (\eta_2-\eta_1)-(\eta_3-\eta_2) = 5-3 = 2 = \delta_2$$
$$k=4:\quad (\eta_4-\eta_3)-(\eta_5-\eta_4) = 3-2 = 1 = \delta_4$$
$$k=5:\quad (\eta_5-\eta_4)-(\eta_6-\eta_5) = 2-0 = 2 = \delta_5$$

Thus, the degrees of the e.d. are (2,2,4,5,5). The PAPSD may be readily constructed as follows:

PAPS diagram:

index:	1	2	3	4	5	6
gap :	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
	•	*	•	*	*	•
		*			*	

and thus the e.d. have degrees {5,5,4,2,2}.

The alternative technique based on the Ferrer's diagram is illustrated next.

Ferrer's diagram:

$$\mathcal{W}_\alpha(F,G) = \{\eta_1-\eta_o=5, \eta_2-\eta_1=5, \eta_3-\eta_2=3, \eta_4-\eta_3=3, \eta_5-\eta_4=2\}$$

The Ferrer's diagram is

$\gamma_1=5$	*	*	*	*	*
$\gamma_2=5$	*	*	*	*	*
$\gamma_3=3$	*	*	*		
$\gamma_4=3$	*	*	*		
$\gamma_5=2$	*	*			
	↓	↓	↓	↓	↓
	5	5	4	2	2

and thus $S_\alpha(F,G) = \{5,5,4,2,2\}$. □

Remark (4.3): The results presented for the determination of the degrees of e.d. at s=α may also be applied for finding the degrees of e.d. at s=∞. The only difference, however, is that the matrices $P^i_\infty(F,G)$ have now to be used instead of the $P^i_\alpha(F,G)$ matrices. This is a mere consequence of the fact that an infinite e.d. of sF-G is a zero-e.d. of $F-\hat{s}G$.

4.4 The structure of nested basis matrices and the maximal generalised nullspace of (F,G) at s=α (∞)

The analysis so far has produced two procedures for the determination of the Segre characteristics of $sF-\hat{s}G$ for all values of α (finite, infinite) for which $|\alpha F-G|=0$. Our attention is focussed now on the geometric aspects of the e.d. set of $sF-\hat{s}G$; those aspects are summarised in the structure of $\mathcal{N}_\alpha^i$ ($\mathcal{N}_\infty^i$), and especially by the properties of the canonical basis matrices N_α^i (N_∞^i) of $\mathcal{N}_\alpha^i$ ($\mathcal{N}_\infty^i$) introduced by Corollary (4.3) and which are of the type

$$N_\alpha^i = \left[\begin{array}{cccccc}
 & \bigcirc & & & 0 & X_i^1 \\
 & & & 0 & X_{i-1}^1 & X_i^2 \\
 & & & & X_{i-1}^2 & X_i^3 \\
 & & 0 & & & \\
 & 0 & X_3^1 & & X_{i-1}^{i-3} & X_i^{i-2} \\
0 & X_2^1 & X_3^2 & & X_{i-1}^{i-2} & X_i^{i-1} \\
X_1^1 & X_2^2 & X_3^3 & & X_{i-1}^{i-1} & X_i^i
\end{array}\right] = \left[\begin{array}{cc|c}
0 & 0 & X_i^1 \\ \hline
 & & X_i^2 \\
 & N_\alpha^{i-1} & \\
 & & X_i^{i-1} \\
 & & X_i^1
\end{array}\right] \tag{4.37}$$

N_α^i is a basis matrix of $\mathcal{N}_\alpha^i$ and N_α^{i-1} is a basis matrix of $\mathcal{N}_\alpha^{i-1}$. Clearly, basis matrices of a similar type may be defined for $\mathcal{N}_\infty^i$. Such matrices will be referred to as <u>nested basis matrices</u> of $\mathcal{N}_\alpha^i$ ($\mathcal{N}_\infty^i$), because of the special properties summarised by eqn(4.37).

The aim of the present section is to provide a detailed analysis of the properties of N_α^i (N_∞^i) nested basis matrices; of special interest are those properties related to the selection of maximal length, linear independent vector chains, where each of these chains characterises an e.d. of (F,G) at $s=\alpha$. The results will be presented for an $s=\alpha\in\mathbb{C}$, whereas the case of $s=\infty$ is similar (the only difference being that the analysis is based on the $P_\infty^i(F,G)$ matrices). It will be shown that the maximal generalised nullspace $\mathcal{M}_\alpha^*$ of (F,G) at $s=\alpha$ may be decomposed as a direct sum of elementary subspaces; an alternative procedure for finding the degrees of the e.d. at $s=\alpha$ will also be given.

We start off by giving a Corollary of Theorem (4.3).

Corollary (4.6): Let $P^{k\perp}_{\alpha} \in \mathbb{C}^{kn \times \eta_k}$ be a right annihilator of $P^{k}_{\alpha}(F,G)$ and let M^{k}_{α} be the submatrix of $P^{k\perp}_{\alpha}$ made up from the last n rows of $P^{k\perp}_{\alpha}$. Then M^{k}_{α} is a basis matrix of the k-th generalised nullspace $\mathcal{M}^{k}_{\alpha}$ of (F,G) at $s=\alpha$.

Proof

From the proof of Theorem (4.3) we have that X is a right annihilator of $P^{k}_{\alpha}(F_w,G_w)$, and that $\hat{X}_k$ has full rank. A basis matrix for $\mathcal{N}^{k}_{\alpha}$ is given by $P^{k\perp}_{\alpha} = \mathrm{diag}\{Q^{-1},\ldots,Q^{-1}\}X$, where Q is the right transformation used in the reduction of (F,G) to its Weierstrass form. Thus, $M^{k}_{\alpha} = Q^{-1}\hat{X}_k$ has full rank and thus it is a basis matrix for $\mathcal{M}^{k}_{\alpha}$.

□

Clearly, the above result also applies to the case of nested basis matrices of $\mathcal{N}^{k}$. Thus, we have the following remark.

Remark (4.4): The submatrix $[X^1_1, X^2_2, \ldots, X^i_i] \in \mathbb{C}^{n \times \eta_i}$ of N^{k}_{α} in (4.37) has rank η_i. The submatrices $X^1_1, X^2_2, \ldots, X^i_i$ have dimensions $n\times\eta_1, n\times(\eta_2-\eta_1), \ldots$ $\ldots, n\times(\eta_i-\eta_{i-1})$ correspondingly and their respective ranks are $\eta_1, \eta_2-\eta_1, \ldots, \eta_i-\eta_{i-1}$. By Corollary (4.5), we also have that $\eta_1 \geq \eta_2-\eta_1 \geq \ldots \geq \eta_i-\eta_{i-1}$.

The problem considered next is the establishment of the relationship between two nested basis matrices of $\mathcal{N}^{i}_{\alpha}$; as a result of this study a canonical decomposition of all nested basis matrices will be given.

Proposition (4.7): Let $N^{i}_{\alpha}, \tilde{N}^{i}_{\alpha}$ be two basis matrices for $\mathcal{N}^{i}_{\alpha}$ of the type (4.37). Then,

$$\tilde{N}^{i}_{\alpha} = N^{i}_{\alpha} P_{i)} \tag{4.38a}$$

where

$$P_{i)} = \left[\begin{array}{c|c|c|c} P^1_1 & P^2_1 & \cdots & P^i_1 \\ \hline & P^2_2 & \cdots & P^i_2 \\ \hline & & \ddots & \\ \hline \bigcirc & & & P^i_i \end{array}\right] \in \mathbb{C}^{\eta_i \times \eta_i} \tag{4.38b}$$

where P^j_j are $(\eta_j-\eta_{j-1})\times(\eta_j-\eta_{j-1})$ nonsingular matrices and $j=1,2,\dots,i$ $(\eta_o=0)$.

Proof

The result is proved by induction. In fact, we prove it for $i=1,2,3$ and the generalisation is rather trivial.

(a) $i=1$: Clearly, $\tilde{X}^1_1, X^1_1$ are basis matrices for $N_r\{G-\alpha F\}$ and thus $\tilde{X}^1_1=X^1_1P^1_1$, $P^1_1\in\mathbb{C}^{\eta_1\times\eta_1}$, $|P^1_1|\neq 0$.

(b) $i=2$: $N^2_\alpha, \tilde{N}^2_\alpha$ as basis matrices of $N_r\{P^2_\alpha(F,G)\}$ are related by

$$\left[\begin{array}{c:c} 0 & \tilde{X}^1_2 \\ \hdashline \tilde{X}^1_1 & \tilde{X}^2_2 \end{array}\right] = \left[\begin{array}{c:c} 0 & X^1_2 \\ \hdashline X^1_1 & X^2_2 \end{array}\right] \left[\begin{array}{c:c} P^1_1 & P^2_1 \\ \hdashline P^1_2 & P^2_2 \end{array}\right]$$

from which

$$\tilde{X}^1_2 = X^1_2P^2_2,\ X^1_2P^1_2 = 0 \tag{4.39a}$$

$$\tilde{X}^1_1 = X^1_1P^1_1+X^2_2P^1_2,\ \tilde{X}^2_2 = X^1_1P^2_1+X^2_2P^2_2 \tag{4.39b}$$

Since $\tilde{X}^1_1, X^1_1$ are basis matrices of $N_r\{G-\alpha F\}$ we have that $\tilde{X}^1_1=X^1_1\tilde{P}^1_1$, where $\tilde{P}^1_1\in\mathbb{C}^{\eta_1\times\eta_1}$, $|\tilde{P}^1_1|\neq 0$ and thus by the first of (4.39b) we have

$$\begin{bmatrix} X^1_1, X^2_2 \end{bmatrix} \begin{bmatrix} \tilde{P}^1_1-P^1_1 \\ P^1_2 \end{bmatrix} = 0 \tag{4.39c}$$

Because $[X^1_1,X^2_2]\in\mathbb{C}^{n\times\eta_2}$, $\eta_2\leq n$, and by Corollary (4.6) has full rank, then $P^1_2=0$ and $\tilde{P}^1_1=P^1_1$ and this completes the proof for $i=2$.

(c) $i=3$: The $N^3_\alpha, \tilde{N}^3_\alpha$ basis matrices of $N_r\{P^3_\alpha(F,G)\}$ are related by

$$\left[\begin{array}{c:c:c} 0 & 0 & \tilde{X}^1_3 \\ \hdashline 0 & \tilde{X}^1_2 & \tilde{X}^2_3 \\ \hdashline \tilde{X}^1_1 & \tilde{X}^2_2 & \tilde{X}^3_3 \end{array}\right] = \left[\begin{array}{c:c:c} 0 & 0 & X^1_3 \\ \hdashline 0 & X^1_2 & X^2_3 \\ \hdashline X^1_1 & X^2_2 & X^3_3 \end{array}\right] \left[\begin{array}{c:c:c} P^1_1 & P^2_1 & P^3_1 \\ \hdashline P^1_2 & P^2_2 & P^3_2 \\ \hdashline P^1_3 & P^2_3 & P^3_3 \end{array}\right]$$

from which

$$X^1_3P^1_3 = 0,\ X^1_3P^2_3 = 0,\ X^1_3P^3_3 = \tilde{X}^1_3 \tag{4.40a}$$

$$X^1_2P^1_2+X^2_3P^1_3 = 0,\ X^1_2P^2_2+X^2_3P^2_3 = \tilde{X}^1_2,\ X^1_2P^3_2+X^2_3P^3_3 = \tilde{X}^2_3 \tag{4.40b}$$

$$X^1_1P^1_1+X^2_2P^1_2+X^3_3P^1_3 = \tilde{X}^1_1,\ X^1_1P^2_1+X^2_2P^2_2+X^3_3P^2_3 = \tilde{X}^2_2,\ X^1_1P^3_1+X^2_2P^3_2+X^3_3P^3_3 = \tilde{X}^3_3 \tag{4.40c}$$

As in step (b), $\tilde{X}^1_1=X^1_1\tilde{P}^1_1$ and thus from the first of (4.40c)

$$\begin{bmatrix} X^1_1, X^2_2, X^3_3 \end{bmatrix} \begin{bmatrix} P^1_1-\tilde{P}^1_1 \\ P^1_2 \\ P^1_3 \end{bmatrix} = 0 \tag{4.40d}$$

$[X^1_1,X^2_2,X^3_3] \in \mathbb{C}^{n\times n_3}$, $n_3 \leq n$, and by Corollary (4.6) has full rank. Thus, $P^1_1=\tilde{P}^1_1$, $P^1_2=0$, $P^1_3=0$.

By step (b) we have that $\tilde{X}^2_2=X^1_1\tilde{P}^2_1+X^2_2\tilde{P}^2_2$; by substituting into the second of (4.40c) equations we have

$$\begin{bmatrix} X^1_1, X^2_2, X^3_3 \end{bmatrix} \begin{bmatrix} P^2_1-\tilde{P}^2_1 \\ P^2_2-\tilde{P}^2_2 \\ P^2_3 \end{bmatrix} = 0 \tag{4.40e}$$

and thus $P^2_1=\tilde{P}^2_1$, $P^2_2=\tilde{P}^2_2$, $P^2_3=0$. The proof is readily completed by induction. □

From the above proof we also have the following Remark.

Remark (4.5): Let $N^i_\alpha, \tilde{N}^i_\alpha$ be two nested bases of $N_r\{P^i_\alpha(F,G)\}$ and let $P_{i)}$ be the transformation of the type (4.38b) for which $\tilde{N}^i=N^iP_{i)}$. If $N^{i+1}_\alpha, \tilde{N}^{i+1}_\alpha$ are basis matrices of $N_r\{P^{i+1}_\alpha(F,G)\}$ obtained from $N^i_\alpha, \tilde{N}^i_\alpha$ by extension as in (4.37), then

$$\tilde{N}^{i+1}_\alpha = N^{i+1}_\alpha P_{i+1)} \tag{4.41a}$$

where

$$P_{i+1)} = \left[\begin{array}{c|c} & P^{i+1}_{1} \\ P_{i)} & \vdots \\ \hline 0 & P^{i+1}_{i+1} \end{array}\right] \tag{4.41b}$$

□

Before we proceed with the study of properties of nested basis matrices of $\mathcal{N}^i_\alpha$ we introduce some useful notation. Let $\{(d_1,\sigma_1),\ldots,(d_\rho,\sigma_\rho);\ d_1<\ldots<d_\rho\}$ be the index set of (F,G) at $s=\alpha$, let $\underline{e}^j_{d_i}$ be the j-th standard basis vector of $\mathbb{R}^{d_i}$, i.e.

$$\underline{e}^j_{d_i} = [\underset{j}{\underleftrightarrow{0,\ldots,0,1}},0\ldots0]^t \in \mathbb{R}^{d_i},\ j=1,2,\ldots d_i \tag{4.42a}$$

and

$$\hat{E}^j_{d_i,\sigma_i} = \left[\begin{array}{cccc} \underline{e}^j_{d_i} & & & \bigcirc \\ & \underline{e}^j_{d_i} & & \\ & & \ddots & \\ \bigcirc & & & \underline{e}^j_{d_i} \end{array}\right] \in \mathbb{R}^{\sigma_i d_i \times \sigma_i}\quad j=1,2,\ldots,d_i \tag{4.42b}$$

If n is the dimension of the regular pencil, then $n=p+\sigma_1 d_1+\ldots+\sigma_\rho d_\rho$ and a trivial expansion of $\hat{E}^j_{d_i,\sigma_i}$ to n dimension is defined by

$$E^j_{d_i,\sigma_i} = \left[\begin{array}{c} 0 \\ \hdashline \vdots \\ \hdashline 0 \\ \hdashline \hat{E}^j_{d_i,\sigma_i} \\ \hdashline 0 \\ \hdashline \vdots \\ \hdashline 0 \end{array}\right]\begin{array}{l} \updownarrow p \\ \\ \updownarrow \sigma_{i-1}d_{i-1} \\ \updownarrow \sigma_i d_i \\ \updownarrow \sigma_{i+1}d_{i+1} \\ \\ \updownarrow \sigma_\rho d_\rho \end{array} \tag{4.42c}$$

$$\xleftrightarrow{\ \ \sigma_i\ \ }$$

With this preliminary notation we may define:

<u>Definition (4.3)</u>: Let $\{(\sigma_1,d_1),\ldots,(\sigma_\rho,d_\rho);d_1<\ldots<d_\rho\}$ be the index set of (F,G) at $s=\alpha$ and let $\phi=\{d_1,d_2,\ldots,d_\rho\}$. For all positive integers $k=1,2,\ldots,d_\rho$ we define as ϕ_k the ordered set of all integers in ϕ which are greater or equal to k, i.e.

$$\begin{aligned} \phi_k &= \{d_i,d_{i+1},\ldots,d_\rho\}, \text{ if } d_{i-1}<k\le d_i \\ \phi_k &= \{d_1,d_2,\ldots,d_\rho\} \quad , \text{ if } k\le d_1 \end{aligned}$$

Assume that $\phi_k=\{d_i,\ldots,d_\rho\}$; then we define by $E^j(\phi_k)$ the matrix

$$E^j(\phi_k) = \left[E^j_{d_i,\sigma_i}, E^j_{d_{i+1},\sigma_{i+1}},\ldots,E^j_{d_\rho,\sigma_\rho} \right] \in \mathbb{R}^{n\times(\sigma_i+\ldots+\sigma_\rho)}$$

□

We may now state the following result:

<u>Proposition (4.8)</u>: Let $\{(d_1,\sigma_1),\ldots,(d_\rho,\sigma_\rho);d_1<\ldots<d_\rho\}$ be the index set of (F,G) at $s=\alpha$ and let $\phi=\{d_1,d_2,\ldots,d_\rho\}$. There exists a nested basis matrix $\hat{N}^k_\alpha$ of $N_r\{P^k_\alpha(F,G)\}$ for all $k=1,2,\ldots,d_\rho$ of the following type

$$\hat{N}^k_\alpha = \text{diag}\{\underbrace{Q,\ldots,Q}_{k}\}E_k \tag{4.43a}$$

where $Q\in\mathbb{C}^{n\times n}$, $|Q|\neq 0$ and

$$E_k = \left[\begin{array}{ccc:c:c} & \bigcirc & & 0 & E^1(\phi_k) \\ \hdashline & & & \iddots & E^2(\phi_k) \\ \hdashline & & 0 & & \vdots \\ \hdashline & 0 & E^1(\phi_3) & \cdots & E^{k-2}(\phi_k) \\ \hdashline 0 & E^1(\phi_2) & E^2(\phi_3) & \cdots & E^{k-1}(\phi_k) \\ \hdashline E^1(\phi_1) & E^2(\phi_2) & E^3(\phi_3) & \cdots & E^k(\phi_k) \end{array}\right] \tag{4.43b}$$

<u>Proof</u>

Let (R^{-1},Q^{-1}) be the pair of transformations which reduce (F,G) to the (4.11) type Weierstrass form (the blocks associated with the e.d. at $s=\alpha$ are the last and they ordered according to the degrees of the e.d.). Then,

$R^{-1}(G-\alpha F)Q^{-1}=G_w-\alpha F_w$, where

$$G_w-\alpha F_w = \mathrm{diag}\{\underset{\xleftrightarrow[p]{}}{T};\underset{\xleftrightarrow[\sigma_1]{}}{H_{d_1},\dots,H_{d_1}};\dots;\underset{\xleftrightarrow[\sigma_\rho]{}}{H_{d_\rho},\dots,H_{d_\rho}}\},\ T\in\mathbb{R}^{\tau\times\tau},|T|\neq 0 \tag{4.44}$$

The result is proved by induction. Thus consider the cases:

(a) <u>k=1</u>: By inspection of (4.44) it is clear that

$$E^1(\phi_1)=\left[E^1_{d_1,\sigma_1},\dots,E^1_{d_\rho,\sigma_2}\right]\in\mathbb{R}^{n\times(\sigma_1+\dots+\sigma_\rho)}$$

is a right annihilator $G_w-\alpha F_w$.

(b) <u>$1<k\leq d_\rho$</u>: From the proof of Theorem (4.3) we have that if $_k$ is a right annihilator of $N_r\{P^k_\alpha(F_w,G_w)\}$, then it may be expressed as

$$X_k=\begin{bmatrix}\tilde{X}^k_1\\ \hline \cdot\\ \cdot\\ \cdot\\ \hline \tilde{X}^k_{k-1}\\ \hline \tilde{X}^k_k\end{bmatrix}=\begin{bmatrix}\begin{bmatrix}0\\ \hat{X}^k_1\end{bmatrix}\\ \vdots\\ \begin{bmatrix}0\\ \hat{X}^k_{k-1}\end{bmatrix}\\ \begin{bmatrix}0\\ \hat{X}^k_k\end{bmatrix}\end{bmatrix}\in\mathbb{R}^{kn\times\eta_k}\ \hat{X}^k_1,\dots\ \hat{X}^k_k\in\mathbb{R}^{(n-\tau)\times\eta_k} \tag{4.45a}$$

where $\hat{X}^k_k$ is a right annihilator of H^k, where

$$H=\mathrm{diag}\{\underset{\xleftrightarrow[\sigma_1]{}}{H_{d_1},\dots,H_{d_1}};\dots;\underset{\xleftrightarrow[\sigma_\rho]{}}{H_{d_\rho},\dots,H_{d_\rho}}\} \tag{4.45b}$$

and $\hat{X}^k_1,\dots\ \hat{X}^k_{k-1}$ are related to $\hat{X}^k_k$ by (4.32e) i.e.

$$\hat{X}^k_{k-1}=H\hat{X}^k_k,\ \hat{X}^k_{k-2}=H^2\hat{X}^k_k,\dots,\hat{X}^k_{k-j}=H^j\hat{X}^k_k,\dots,\ \hat{X}^k_1=H^{k-1}\hat{X}^k_k \tag{4.45c}$$

If we define by $\tilde{H}$ the matrix

$$\tilde{H}=\mathrm{diag}\{I_p;H\} \tag{4.46a}$$

then by (4.45c) we have that the matrices $\tilde{X}^k_{k-j}$, h=1,...,k-1 are related to $\tilde{X}^k_k$ by

$$\tilde{X}^k_{k-1} = \tilde{H}\tilde{X}^k_k,\ \tilde{X}^k_{k-2} = \tilde{H}^2\tilde{X}^k_k, \ldots, \tilde{X}^k_{k-j} = \tilde{H}^j\tilde{X}^k_k, \ldots, \tilde{X}^k_1 = \tilde{H}^{k-1}\tilde{X}^k_k \tag{4.56b}$$

The problem of defining the structure of X_k is then reduced to a problem of computing the structure of $\tilde{X}^k_k$. Since $\tilde{X}^k_k$ is a right annihilator of H^k, it is readily verified that a choice for $\tilde{X}^k_k$ is given by

$$\tilde{X}^k_k = [E^1(\phi_1), E^2(\phi_2), E^3(\phi_3), \ldots, E^k(\phi_k)] \tag{4.46c}$$

which is a right annihilator of $\tilde{H}^k$. From the special structure of the $E^i(\phi_i)$ matrices and the compatible special structure of $\tilde{H}$ we may easily verify the following property

$$\tilde{H}^p E^i(\phi_i) = \begin{cases} 0, & \text{if } p \geq i \\ E^{i-p}(\phi_i), & \text{if } 0<p<i \end{cases} \tag{4.47}$$

From (4.47) and (4.46b) it then follows that E_k is a right annihilator of $P^k_\alpha(F_w, G_w)$, when $G_w - \alpha F_w$ is expressed in the (4.44) form. However, with an appropriate choice of the $(\bar{R}^{-1}, \bar{Q}^{-1})$ pair, $G_w - \alpha F_w$ may always be expressed as in (4.44). It is obvious that if E_k is a right annihilator of $P^k_\alpha(F_w, G_w)$, then $\text{diag}\{\underbrace{Q, \ldots, Q}_{k}\}E_k$ is a right annihilator of $P^k_\alpha(F, G)$.

□

The special structure of the $E^i(\phi_j)$ matrices clearly implies the following property.

Remark (4.6): If $A \in \mathbb{R}^{m\times n}$, $A' \in \mathbb{R}^{m\times n'}$, $n' \leq n$, then we shall write $A' \subseteq^c A$, if the columns of A' is a subset of the columns of A. For the matrices $E^i(\phi_j)$ in E_k we have that

$$\begin{aligned} &E^1(\phi_k) \subseteq^c E^1(\phi_{k-1}) \subseteq^c \cdots \subseteq^c E^1(\phi_3) \subseteq^c E^1(\phi_2) \subseteq^c E^1(\phi_1) \\ &E^2(\phi_k) \subseteq^c E^2(\phi_{k-1}) \subseteq^c \cdots \subseteq^c E^2(\phi_3) \subseteq^c E^2(\phi_2) \\ &\qquad \vdots \qquad\qquad \vdots \\ &E^{k-2}(\phi_n) \subseteq^c E^{k-2}(\phi_{n-1}) \subseteq^c E^{k-2}(\phi_{n-2}) \\ &E^{k-1}(\phi_n) \subseteq^c E^{k-2}(\phi_{n-1}) \end{aligned} \tag{4.48}$$

□

By combining Propositions (4.8) and (4.7) we have the following characterisation of nested basis matrices of $\mathcal{N}_r\{P_\alpha^i(F,G)\}$.

<u>Theorem (4.4)</u>: Every nested basis matrix N_α^k of $\mathcal{N}_r\{P_\alpha^k(F,G)\}$ may be expressed as

$$N_\alpha^k = \left[\begin{array}{cc:c:c:c:c} \bigcirc & & & & 0 & X_k^1 \\ \hdashline & & & & X_{k-1}^1 & X_k^2 \\ \hdashline & & 0 & \cdot^{\cdot^{\cdot}} & \vdots & \vdots \\ \hdashline & 0 & X_3^1 & \cdots & X_{k-1}^{k-3} & X_k^{k-2} \\ \hdashline 0 & X_2^1 & X_3^2 & \cdots & X_{k-1}^{k-2} & X_k^{k-1} \\ \hdashline X_2^1 & X_2^2 & X_3^3 & \cdots & X_{k-1}^{k-1} & X_k^k \end{array}\right] = \mathrm{diag}\underbrace{\{Q,\ldots,Q\}}_{k} E_k P_{k)} \tag{4.49}$$

where $Q\in\mathbb{C}^{n\times n}$, $|Q|\neq 0$, $P_{k)}\in\mathbb{C}^{\eta_k\times\eta_k}$, $|P_{k)}|\neq 0$ and with a structure defined by (4.38b) and E_k the matrix defined by (4.43b). □

<u>Corollary (4.7)</u>: Let N_α^k be a nested basis matrix of $\mathcal{N}_\alpha^k$. For every i, $i=1,2,\ldots,k$ and for every vector $\underline{\alpha}\in\mathbb{C}^{(\eta_i-\eta_{i-1})}$, $\underline{\alpha}\neq\underline{0}$. the set of vectors $\{X_i^i\underline{\alpha}, X_i^{i-1}\underline{\alpha},\ldots,X_i^1\underline{\alpha}\}$ are linearly independent.

<u>Proof</u>

Assume that the set of vectors $\{X_i^i\underline{\alpha},\ldots,X_i^1\underline{\alpha}\}$ is linearly dependent. Then, there exist c_j, $j=1,2,\ldots,i$ not all of them zero such that

$$c_i X_i^i\underline{\alpha}+c_{i-1}X_i^{i-1}\underline{\alpha}+\ldots+c_1X_i^1\underline{\alpha} = 0 \tag{4.50a}$$

By Theorem (4.4) we have that $X_i^i,\ldots,X_i^1$ may be expressed as

$$\begin{aligned} X_i^1 &= QE^1(\phi_i)P_i^i \\ X_i^2 &= Q\{E^1(\phi_{i-1})P_{i-1}^i+E^2(\phi_i)P_i^i\} \\ \\ X_i^{i-1} &= Q\{E^1(\phi_2)P_2^i+E^2(\phi_3)P_3^i+\ldots+E^{i-2}(\phi_{i-1})P_{i-1}^i+E^{i-1}(\phi_i)P_i^i\} \\ X_i^i &= Q\{E^1(\phi_1)P_1^i+E^2(\phi_2)P_2^i+\ldots+E^{i-1}(\phi_{i-1})P_{i-1}^i+E^i(\phi_i)P_i^i\} \end{aligned} \tag{4.50b}$$

If we define $P^i_i\underline{\alpha}=\underline{\alpha}_i$, $P^i_{i-1}\underline{\alpha}=\underline{\alpha}_{i-1},\dots,$ $P^i_1\underline{\alpha}=\underline{\alpha}_1$, then by (4.50b) we have that (4.50a) may be expressed as

$$\begin{aligned}Q\{&c_1E^1(\phi_i)\underline{\alpha}_i+c_2(E^1(\phi_{i-1})\underline{\alpha}_{i-1}+E^2(\phi_i)\underline{\alpha}_i)+\dots+\\ &+c_{i-1}(E^1(\phi_2)\underline{\alpha}_2+E^2(\phi_3)\underline{\alpha}_3+\dots+E^{i-2}(\phi_{i-1})\underline{\alpha}_{i-1}+E^{i-1}(\phi_i)\underline{\alpha}_i)+\\ &+c_i(E^1(\phi_1)\underline{\alpha}_1+E^2(\phi_2)\underline{\alpha}_2+\dots+E^{i-1}(\phi_{i-1})\underline{\alpha}_{i-1}+E^i(\phi_i)\underline{\alpha}_i)\}=0\end{aligned} \tag{4.50c}$$

Because Q is nonsingular (4.50c) yields

$$\begin{aligned}-c_iE^i(\phi_i)\underline{\alpha}_i &= c_1E^1(\phi_i)\underline{\alpha}_i+c_2(E^1(\phi_{i-1})\underline{\alpha}_{i-1}+E^2(\phi_i)\underline{\alpha}_i)+\dots+\\ &\quad+c_i(E^1(\phi_1)\underline{\alpha}_1+E^2(\phi_2)\underline{\alpha}_2+\dots+E^{i-1}(\phi_{i-1})\underline{\alpha}_{i-1})\end{aligned} \tag{4.50d}$$

By Remark (4.6), it is clear that the right hand side of (4.50d) is a vector in $sp\{[E^1(\phi_1),E^2(\phi_2),\dots,E^{i-1}(\phi_{i-1})]\}$; however, since the left hand side is a vector in $sp\{E^i(\phi_i)\}$ and that $sp\{[E^1(\phi_1),\dots,E^{i-1}(\phi_{i-1})]\}\cap$ $\cap sp\{E^i(\phi_i)\}=\{0\}$ we have that $c_i\underline{\alpha}_i=0$. If $\underline{\alpha}_i=0$, then $P^i_i\underline{\alpha}=0$ which implies that $\underline{\alpha}=0$, since $|P^i_i|\neq0$. Thus, it is shown that $c_i=0$. Set $c_i=0$ in (4.50c) and by repeating the same arguments it follows that $c_{i-1}=0$ etc. Thus, eventually $c_1=c_2=\dots=c_i=0$, which contradicts the linear dependence assumption. □

If we choose $\underline{\alpha}=[0,\dots,0,1,0,\dots0]^t\in\mathbb{R}^{(n_i-n_{i-1})}$, then we select a column of N^k_α and Corollary (4.7) yields the following important property.

<u>Remark (4.7)</u>: The set of nonzero vectors obtained by partitioning every column of a nested basis matrix N^k_α, according to the natural partitioning of N^k_α (as in (4.49)) is linearly independent.

The vector chain $\{X^i_i\underline{\alpha},X^{i-1}_i\underline{\alpha},\dots,X^1_i\underline{\alpha}\}$ is by Corollary (4.7) linearly independent, it corresponds to the i-th column block of N^k_α and it is parametrised by $\underline{\alpha}$; such vector chains we shall denote then by $S(i,\underline{\alpha})$. The properties of $S(i,\underline{\alpha})$ chains are examined next.

Corollary (4.8): Let N_{α}^{k} be a nested basis matrix of $\mathcal{N}_r\{P_{\alpha}^{k}(F,G)\}$ and let $S(i,\underline{\alpha}_j)$ be the vector chains associated with the i-th column block of N_{α}^{k} and parametrised by the $\underline{\alpha}_j$ vectors. The set of vectors $S(i)=\{S(i,\underline{\alpha}_1);\dots;\dots;S(i,\underline{\alpha}_\nu)\}$ is linearly independent if and only if either of the following equivalent conditions hold true:

(i) the set $\{\underline{\alpha}_1,\underline{\alpha}_2,\dots,\underline{\alpha}_\nu\}$ is independent.

(ii) the set $\{X_i^1\underline{\alpha}_1,X_i^1\underline{\alpha}_2,\dots,X_i^1\underline{\alpha}_\nu\}$ is independent.

Proof

Let us assume that the set S(i) is linearly dependent. Then, there exist constants c_k^j, $k=1,\dots,\nu$, $j=1,2,\dots,i$, not all of them zero such that

$$\Big(\sum_{j=1}^{i} c_1^j X_i^j\Big)\underline{\alpha}_1+\Big(\sum_{j=1}^{i} c_2^j X_i^j\Big)\underline{\alpha}_2+\dots+\Big(\sum_{j=1}^{i} c_\nu^j X_i^j\Big)\underline{\alpha}_\nu = 0$$

This may be rewritten as

$$X_i^i\Big(\sum_{k=1}^{\nu} c_k^i\underline{\alpha}_k\Big)+X_i^{i-1}\Big(\sum_{k=1}^{\nu} c_k^{i-1}\underline{\alpha}_k\Big)+\dots+X_i^2\Big(\sum_{k=1}^{\nu} c_k^2\underline{\alpha}_k\Big)+X_i^1\Big(\sum_{k=1}^{\nu} c_k^1\underline{\alpha}_k\Big) = 0 \qquad (4.51a)$$

Let us now denote by

$$\underline{\beta}_j = \sum_{k=1}^{\nu} c_k^j\underline{\alpha}_k \text{ and } P_r^i\underline{\beta}_j = \underline{\beta}_{r,j},\ r,j=1,2,\dots,i \qquad (4.51b)$$

By (4.51b) and (4.50b), condition (4.51a) is equivalent to

$$\begin{aligned}&\{E^1(\phi_i)\underline{\beta}_{i,1}\}+\{E^1(\phi_{i-1})\underline{\beta}_{i-1,2}+E^2(\phi_i)\underline{\beta}_{i,2}\}+\dots\\ &\quad+\{E^1(\phi_2)\underline{\beta}_{2,i-1}+E^2(\phi_3)\underline{\beta}_{3,i-1}+\dots+E^{i-1}(\phi_i)\underline{\beta}_{i,i-1}\}+\\ &\quad+\{E^1(\phi_1)\underline{\beta}_{1,i}+E^2(\phi_2)\underline{\beta}_{2,i}+\dots+E^{i-1}(\phi_{i-1})\underline{\beta}_{i-1,i}+E^i(\phi_i)\underline{\beta}_{i,i}\} = 0 \qquad (4.51c)\end{aligned}$$

The above condition is similar in nature to (4.50d); thus, by using similar arguments (based on the properties of $E^j(\phi_k)$ matrices) it follows that

$$\underline{\beta}_{i,i} = 0 = P_i^i\underline{\beta}_i \implies \underline{\beta}_i = 0$$

Repeating the arguments on the reduced equation (4.51c) (after setting $\underline{\beta}_i=0$) we have that

$$\underline{\beta}_j = \sum_{k=1}^{\nu} c_k^j \underline{\alpha}_k = 0, \; j=1,2,\dots,i \tag{4.51d}$$

If the vectors $\{\underline{\alpha}_1,\dots,\underline{\alpha}_\nu\}$ are linearly independent, then (4.51d) conditions imply that $c_k^j=0$ for all $k=1,2,\dots\nu$, $j=1,2,\dots i$ and thus the assumption that S(i) is dependent leads to a contradiction. If the vectors $\{\underline{\alpha}_1,\dots,\underline{\alpha}_\nu\}$ are dependent, then at least one of (4.51d) may be satisfied with nonzero constants and thus S(i) is dependent.

To prove the equivalence of parts (i) and (ii) we notice that if $\underline{\alpha}_j$ are linearly dependent, then

$$\sum_{j=1}^{\nu} c_j \underline{\alpha}_j = 0 \tag{4.52a}$$

By the first of (4.50b), $X_i^1=QE^1(\phi_i)P_i^i$ and thus X

$$X_i^1 \sum_{j=1}^{\nu} c_j \underline{\alpha}_j = 0 = \sum_{j=1}^{\nu} c_j X_i^1 \underline{\alpha}_j \tag{4.52b}$$

If the set $\{\underline{\alpha}_j, j\in\underline{\nu}\}$ is independent and $\{X_i^1\underline{\alpha}_i, i\in\underline{\nu}\}$ dependent then (4.52b) and the fact that $\mathcal{N}_r\{QE^1(\phi_i)P_i^i\}=\{0\}$, yields that $\{\underline{\alpha}_j, j\in\underline{\nu}\}$ is dependent, which contradicts our assumption.

□

If we select $\underline{\alpha}_j=\underline{e}_j$, $i=1,2,\dots,\eta_i-\eta_{i-1}$, where $\underline{e}_j$ are the standard basis vectors of $\mathbb{R}^{(\eta_i-\eta_{i-1})}$, then every vector chain $S(i,\underline{e}_j)=\{X_i^i\underline{e}_j, X_i^{i-1}\underline{e}_j,\dots \dots, X_i^1\underline{e}_j\}$ is independent and thus they form a basis for the vector space

$$S_j^i = sp\{X_i^i\underline{e}_j,\dots,X_i^1\underline{e}_j\} \tag{4.53}$$

There exist $\eta_i-\eta_{i-1}$ subspaces S_j^i, $j=1,2,\dots,\eta_i-\eta_{i-1}$ which characterise the i-th column block of the given N_α^k nested basis matrix. By the way nested basis matrices are constructed, it is clear that the i-th dimensional subspaces S_j^i are characteristic of the N_α^τ nested basis, where τ is the index of annihilation, and not just of the given N_α^k; such subspaces will be referred to as <u>i-th order characteristic spaces</u> of (F,G) at $s=\alpha$. By Corollary (4.8) we have:

<u>Remark (4.8)</u>: The set $\{S_j^i;\ j=1,2,\ldots,\eta_i-\eta_{i-1}\}$ of i-th order characteristic spaces of (F,G) at $s=\alpha$ are linearly independent for any given i, $i=1,2,\ldots,\tau$, where τ is the index of annihilation of (F,G) at $s=\alpha$.

The relationship of the set of S_j^i subspaces for all i,j to the maximal generalised nullspace M_α^* of (F,G) at $s=\alpha$ is established by the following result.

<u>Proposition (4.9)</u>: If M_α^* is the maximal generalised nullspace of (F,G) at $s=\alpha$ and $\{S_j^i,\ j=1,2,\ldots,\eta_i-\eta_{i-1}\}$ the i-th order characteristic spaces of (F,G) at $s=\alpha$, then

$$M_\alpha^* = \sum_{i=1}^{\tau} \{ \sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i \} \tag{4.54}$$

<u>Proof</u>

By definition of the S_j^i we have that

$$\sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i = sp\{X_i^i\}+sp\{X_i^{i-1}\}+\ldots+sp\{X_i^1\}$$

and thus

$$\sum_{i=1}^{\tau} \{ \sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i \} = sp\{X_1^1\}+sp\{X_2^2\}+\ldots+sp\{X_i^i\}+\ldots+sp\{X_\tau^\tau\}$$
$$+sp\{X_2^1\}+\ldots+sp\{X_i^{i-1}\}+\ldots+sp\{X_\tau^{\tau-1}\}+\ldots$$

from which since $M_\alpha^*=sp\{X_1^1\}+\ldots+sp\{X_\tau^\tau\}$ we have that

$$M_\alpha^* \subseteq \sum_{i=1}^{\tau} \{ \sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i \} \tag{4.55a}$$

By eqn(4.50b) we have that

$$\begin{aligned}
&\mathrm{sp}\{X_i^1\} \subseteq \mathrm{sp}\{QE^1(\phi_i)\} \\
&\mathrm{sp}\{X_i^2\} \subseteq \mathrm{sp}\{Q[E^1(\phi_{i-1}),E^2(\phi_i)]\} \\
&\mathrm{sp}\{X_i^{i-1}\} \subseteq \mathrm{sp}\{Q[E^1(\phi_2),E^2(\phi_3),\dots,E^{i-1}(\phi_i)]\} \\
&\mathrm{sp}\{X_i^i\} \subseteq \mathrm{sp}\{Q[E^1(\phi_1),E^2(\phi_2),\dots,E^{i-1}(\phi_{i-1}),E^i(\phi_i)]\}
\end{aligned} \tag{4.55b}$$

By Remark (4.6) and conditions (4.55b) it is clear that

$$\mathrm{sp}\{X_i^j\} \subseteq \mathrm{sp}\{Q[E^1(\phi_1),E^2(\phi_2),\dots,E^j(\phi_j)]\},\ j=1,2,\dots,i \tag{4.55c}$$

and thus

$$\begin{aligned}
\sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i = \mathrm{sp}\{X_i^1\}+\dots+\mathrm{sp}\{X_i^{i-1}\}+\mathrm{sp}\{X_i^i\} &\subseteq \mathrm{sp}\{Q[E^1(\phi_1)]\}+\dots \\
\dots+\mathrm{sp}\{Q[E^1(\phi_1),\dots,E^i(\phi_i)]\} &\subseteq \mathrm{sp}\{Q[E^1(\phi_1),\dots,E^i(\phi_i)]\}
\end{aligned} \tag{4.55d}$$

Clearly then

$$\begin{aligned}
\sum_{i=1}^{\tau} \{ \sum_{j=1}^{\eta_i-\eta_{i-1}} S_j^i \} &\subseteq \mathrm{sp}\{Q[E^1(\phi_1)]\}+\dots+\mathrm{sp}\{Q[E^1(\phi_1),\dots,E^\tau(\phi_\tau)]\} \\
&\subseteq \mathrm{sp}\{Q[E^1(\phi_1),\dots,E^\tau(\phi_\tau)]\} = M_\alpha^*
\end{aligned} \tag{4.55e}$$

By (4.55a) and (4.55e) condition (4.54) is established. □

Note that a vector chain $S(i,\underline{\alpha})=\{X_i^i\underline{\alpha},X_i^{i-1}\underline{\alpha},\dots,X_i^1\underline{\alpha}\}=\{\underline{u}_i^i,\underline{u}_i^{i-1},\dots,\underline{u}_i^1\}$ is an independent chain of vectors (not necessarily maximal) which satisfies conditions (4.4) with a first vector $\underline{u}_i^i=X_i^i\underline{\alpha}$. The above result then implies the following remark.

Remark (4.9): All generalised eigenvector chains $S(i,\underline{\alpha})$ satisfying conditions (4.4) belong to M_α^* and M_α^* is spanned by the set of all chains $S(i,\underline{\alpha})$ defined by any nested basis matrix.

Clearly, this remark generalises the well-known property for the generalised eigenvector chains of $A \in \mathbb{R}^{n\times n}$.

The problem considered next is the investigation of the conditions under which various vector chains of the $S(i,\underline{\alpha})$ type (corresponding to different i and $\underline{\alpha}$ vectors) are linearly independent. The following lemma is useful in this investigation.

Lemma (4.2): Let $\Gamma_j \in \mathbb{R}^{n\times\eta_k}$ be the submatrices of E_k defined by

$$\Gamma_j = [\underset{j-1}{\underleftrightarrow{0,0,\dots,0}},E_j^1,E_{j+1}^2,\dots,E_k^{k+1-j}],\ j=1,2,\dots,k \tag{4.56a}$$

where E_r^p denotes in short $E^p(\phi_r)$, and let $\underline{z}_j=[\underline{z}_{j,1}^t,\underline{z}_{j,2}^t,\dots,\underline{z}_{j,k}^t]^t \in \mathbb{C}^{\eta_k}$, $j=1,2,\dots\underline{k}$ be arbitrary vectors partitioned according to the block partitioning of Γ_j in (4.56a). The matrix equation

$$\Gamma_1\underline{z}_1+\Gamma_2\underline{z}_2+\dots+\Gamma_k\underline{z}_k = 0 \tag{4.56b}$$

implies the following equivalent conditions.

(i)

$$E_i^i\underline{z}_{1,i}+E_{i+1}^i\underline{z}_{2,i+1}+\dots+E_k^i\underline{z}_{k-i+1,k} = 0 \tag{4.56c}$$

for all $i=1,2,\dots,k$.

(ii) If we trivially expand the vectors $\underline{z}_{2,i+1},\dots,\underline{z}_{k-i+1,k}$ to vectors of $\mathbb{C}^{\eta_i-\eta_{i-1}}$ by adding zeros on top of them, then

$$\underline{z}_{1,i}+\begin{bmatrix}\underline{0}\\ \underline{z}_{2,i+1}\end{bmatrix}+\begin{bmatrix}\underline{0}\\ \underline{0}\\ \underline{z}_{3,i+2}\end{bmatrix}+\dots+\begin{bmatrix}\underline{0}\\ \underline{0}\\ \vdots\\ \underline{0}\\ \underline{z}_{k-i+1,k}\end{bmatrix} = 0 \tag{4.56d}$$

for all $i=1,2,\dots,k$.

Proof

By using the partitioned forms of $\underline{z}_j$ and Γ_j, it may be readily verified that (4.56c) yields

$$\underline{\gamma}_1+\dots+\underline{\gamma}_i+\dots+\underline{\gamma}_k = 0 \tag{4.57a}$$

where the $\underline{\gamma}_i$ vectors are given by

$$\underline{\gamma}_i = E_i^i\underline{z}_{1,i} + E_{i+1}^i\underline{z}_{2,i+1} + \ldots + E_k^i\underline{z}_{k-i+1,k} \tag{4.57b}$$

By Remark (4.6), it is clear that $\underline{\gamma}_i \in \mathrm{sp}\{E_i\}$ and given that the set of subspaces $\{\mathrm{sp}\{E_1^1\},\ldots,\mathrm{sp}\{E_i^i\}\}$ are linearly independent for all $i=1,2,\ldots,k$ it follows that $\underline{\gamma}_1=\underline{\gamma}_2=\ldots=\underline{\gamma}_k=0$, which proves part (i) of the lemma.

By Remark (4.6) and the definition of E_j^i, $j=i,\ldots,k$ we have that E_i^i may be partitioned as

$$\begin{aligned} E_i^i &= [E_{i,i+1}^i, E_{i+1,i+2}^i, \ldots, E_{k-1,k}^i E_k^i] \\ &= [E_{i,i+1}^i, E_{i+1,i+2}^i, \ldots, E_{k-2,k-1}^i, E_{k-1}^i] \\ &= \ldots = [E_{i,i+1}^i, E_{i+1,i+2}^i, E_{i+2}^i] = [E_{i,i+1}^i, E_{i+1}^i] \end{aligned} \tag{4.57c}$$

With this in mind, eqn(4.56c) may be expressed as

$$E_i^i\underline{z}_{1,i} + \underbrace{[E_{i,i+1}^i, E_{i+1}^i]}_{\triangleq\, E_i^i}\begin{bmatrix}\underline{0}\\ \underline{z}_{2,i+1}\end{bmatrix} + \ldots + \underbrace{[E_{i,i+1}^i,\ldots,E_{k-1,k}^i E_k^i]}_{\triangleq\, E_i^i}\begin{bmatrix}\underline{0}\\ \vdots\\ \underline{0}\\ \underline{z}_{k-i+1,k}\end{bmatrix} = 0$$

given that $\mathcal{N}_r\{E_i^i\}=\{0\}$, (4.56d) follows. □

The independence of the $S(i,\underline{\alpha})$ vector chains, for various indices i and vectors $\underline{\alpha}$ is characterised by the following result.

<u>Proposition (4.10)</u>: Let N_α^k be a nested basis matrix of $\mathcal{N}_\alpha^k$, $\{i_1,\ldots,i_\nu\}$ be a set of indices taking values from $\{1,2,\ldots,k\}$ and let $\underline{\alpha}_1,\ldots,\underline{\alpha}_\nu$ be vectors of $\mathbb{C}^{(\eta_{i_1}-\eta_{i_1-1})},\ldots,\mathbb{C}^{(\eta_{i_\nu}-\eta_{i_{\nu-1}})}$ respectively.

The set of vector chains $S(i_1,\ldots,i_n)=\{S(i_1,\underline{\alpha}_1);\ldots;S(i_\nu,\underline{\alpha}_\nu)\}$ are linearly independent if and only if the vectors

$$\{X_{i_1}^1\underline{\alpha}_1, X_{i_2}^1\underline{\alpha}_2,\ldots,X_{i_\nu}^1\underline{\alpha}_\nu\}$$

are linearly independent.

Proof

The necessity of the proposition (the set $S(i_1,\dots,i_n)$ independent, then $\{X_{i_1}^1\underline{\alpha}_1,\dots,X_{i\nu}^1\underline{\alpha}_\nu\}$ is independent) is obvious.

Sufficiency: Assume that the set of vectors $\{X_{i_1}^1\underline{\alpha}_1,\dots,X_{i\nu}^1\underline{\alpha}_\nu\}$ are linearly independent and the vector chains $S(i_1,\dots,i_\nu)$ are dependent and that $i_1\leq i_2\leq\dots\leq i_\nu$. Then there exist coefficients $c_{p,j}$, $p=1,\dots,i_j$, $j=i_1,\dots,i_\nu$, not all of them zero, such that

$$\sum_{p=1}^{i_1} c_{p,i_1} X_{i_1}^{i_1-p+1}\underline{\alpha}_1+\sum_{p=1}^{i_2} c_{p,i_2} X_{i_2}^{i_2-p+1}\underline{\alpha}_2+\dots+\sum_{p=1}^{i_\nu} c_{p,i\nu} X_{i_\nu}^{i_\nu-p+1}\underline{\alpha}_\nu=0 \tag{4.58a}$$

By Theorem (4.4), the vectors in the chains may be computed by taking the $\underline{\alpha}_j$ linear combination of the columns of the i_j-th column block of N_α^k and then partitioning according to the natural partitioning of the N_α^k. Thus, for the $S(i_j,\underline{\alpha}_j)$ chain we have

$$\begin{array}{c}k-i_j\\ \text{zero blocks}\end{array}\updownarrow\begin{bmatrix}0\\ \vdots\\ 0\\ X_{i_j}^1\\ X_{i_j}^2\\ \vdots\\ X_{i_j}^{i_j-1}\\ X_{i_j}^{i_j}\end{bmatrix}\underline{\alpha}_j=\mathrm{diag}\{\underbrace{Q,\dots,Q}_{k}\}E_k\underline{p}(i_j,\underline{\alpha}_j) \tag{4.58b}$$

where $\underline{p}(i_j,\underline{\alpha}_j)$ is the $\underline{\alpha}_j$ linear combination of the columns of the i_j column block of $P_{k)}$, i.e.

$$\underline{P}(i_j,\underline{\alpha}_j) = \begin{bmatrix} P_1^{i_j} \\ P_2^{i_j} \\ \vdots \\ P_{i_j}^{i_j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \underline{\alpha}_j = \begin{bmatrix} \underline{u}_{j,1} \\ \underline{u}_{j,2} \\ \vdots \\ \underline{u}_{j,i_j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \begin{matrix} k-i_j \\ \text{zero blocks} \end{matrix} \tag{4.58c}$$

Because $Q \in \mathbb{C}^{n\times n}$, $|Q|\neq 0$, in the investigation of (4.58a), the transformation Q in (4.58b) may be ignored and we can work with the subvectors of $E_k\underline{P}(i_j,\underline{\alpha}_j)$. Using the Γ_ℓ matrices, as defined in (4.56a), we have that

$$E_k\underline{P}(i_j,\underline{\alpha}_j) = \begin{bmatrix} \Gamma_k\underline{P}(i_j,\underline{\alpha}_j) \\ \vdots \\ \Gamma_{i_j+1}\underline{P}(i_j,\underline{\alpha}_j) \\ \Gamma_{i_j}\underline{P}(i_j,\underline{\alpha}_j) \\ \vdots \\ \Gamma_1\underline{P}(i_j,\underline{\alpha}_j) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ Q^{-1}X_{i_j}^1 \\ \vdots \\ Q^{-1}X_{i_j}^{i_j} \end{bmatrix} \underline{\alpha}_j \tag{4.58d}$$

To facilitate the analysis we expand trivially the summations in (4.58a) by including the zero vectors $\Gamma_k\underline{P}(i_j,\underline{\alpha}_j),\ldots,\Gamma_{i_j+1}\underline{P}(i_j,\underline{\alpha}_j)$ with zero coefficients as

$$\sum_{j=1}^{k} c_{j,i_1}\Gamma_j\underline{P}(i_1,\underline{\alpha}_1)+\ldots+\sum_{j=1}^{k} c_{j,i_\nu}\Gamma_j\underline{P}(i_\nu,\underline{\alpha}_\nu) = 0 \tag{4.58e}$$

where $c_{j,i_1}=0$ for $j>i_1,\ldots,c_{j,i_\nu}=0$ for $j>i_\nu$. If we define

$$\underline{z}_\ell = c_{\ell,i_1}\underline{P}(i_1,\underline{\alpha}_1)+\ldots+c_{\ell,i_\nu}\underline{P}(i_\nu,\underline{\alpha}_\nu) \quad \ell=1,\ldots,k \tag{4.58f}$$

then (4.58e) may be rewritten as

$$\Gamma_1\underline{z}_1+\Gamma_2\underline{z}_2+\ldots+\Gamma_k\underline{z}_k = \underline{0} \tag{4.59a}$$

By partitioning $\underline{z}_\ell$ according to the partitioning of Γ_ℓ as $\underline{z}_\ell=[\underline{z}_{\ell,1}^t,\underline{z}_{\ell,2}^t,\ldots,\underline{z}_{\ell,k}^t]^t$, then by Lemma (4.2) we have

$$\underline{z}_{1,i}+\begin{bmatrix}\underline{0}\\ \underline{z}_{2,i+1}\end{bmatrix}+\begin{bmatrix}\underline{0}\\ \underline{0}\\ \underline{z}_{3,i+2}\end{bmatrix}+\ldots+\begin{bmatrix}\underline{0}\\ \underline{0}\\ \vdots\\ \underline{0}\\ \underline{z}_{k-i+1,k}\end{bmatrix}=\underline{0} \qquad (4.59b)$$

for all i=1,2,...,k. By (4.58c) and (4.58f) we have that

$$\underset{i=1,2,\ldots,k}{\underline{z}_\ell=}\begin{bmatrix}c_{\ell,i_1}\underline{u}_{1,1}+c_{\ell,i_2}\underline{u}_{2,1}+\ldots+c_{\ell,i_\nu}\underline{u}_{\nu,1}\\ c_{\ell,i_1}\underline{u}_{1,2}+c_{\ell,i_2}\underline{u}_{2,2}+\ldots+c_{\ell,i_\nu}\underline{u}_{\nu,2}\\ \vdots\\ c_{\ell,i_1}\underline{u}_{1,k}+c_{\ell,i_2}\underline{u}_{2,k}+\ldots+c_{\ell,i_\nu}\underline{u}_{\nu,k}\end{bmatrix}=\begin{bmatrix}\underline{z}_{\ell,1}\\ \underline{z}_{\ell,2}\\ \vdots\\ \underline{z}_{\ell,k}\end{bmatrix} \qquad (4.59c)$$

where $\underline{u}_{j,f}=\underline{0}$ for all $f>i_j$, $j=1,2,\ldots\nu$. The original problem is thus reduced to proving that conditions (4.59b),(4.59c) have as the only solution $c_{\ell,p}=0$ for all $\ell=1,2,\ldots,k$, $p=i_1,\ldots,i_\nu$, if the vectors $\{X_{i_1}^1\underline{\alpha}_1,\ldots,X_{i_\nu}^1\underline{\alpha}_\nu\}$ are independent.

Note that from (4.58d) we have that

$$X_{i_1}^1\underline{\alpha}_1=E_{i_1}^1P_{i_1}^{i_1}\underline{\alpha}_1=E_{i_1}^1\underline{u}_{1,i_1},\ldots X_{i_\nu}^1\underline{\alpha}_\nu=E_{i_\nu}^1P_{i_\nu}^{i_\nu}\underline{\alpha}_\nu=E_{i_\nu}^1\underline{u}_{\nu,i_\nu} \qquad (4.59d)$$

and by (4.57c) we may write

$$E_1^1=\left[E_{1,2}^1,\ldots,E_{i_1-1,i_1}^1,E_{i_1}^1\right]=\left[E_{1,2}^1,\ldots,E_{i_1-1,i_1}^1,E_{i_1,i_1+1}^1,\ldots,E_{i_\nu-1,i_\nu}^1,E_{i_\nu}^1\right]$$

Thus, the vectors $X_{i_1}^1\underline{\alpha}_1,\ldots,X_{i_\nu}^1\underline{\alpha}_\nu$ may be expressed as

$$X_{i_1}^1\underline{\alpha}_1=E_1^1\begin{bmatrix}\underline{0}\\ \underline{u}_{1,i_i}\end{bmatrix}=E_1^1\tilde{\underline{u}}_{1,i_1},\ldots,X_{i_\nu}^1\underline{\alpha}_\nu=E_1^1\begin{bmatrix}\underline{0}\\ \underline{u}_{\nu,i_\nu}\end{bmatrix}=E_1^1\tilde{\underline{u}}_{\nu,i_\nu} \qquad (4.59e)$$

where $\underline{u}_{1,i_1},\ldots,\underline{u}_{\nu,i_\nu}$ are $(\eta_{i_1}-\eta_{i_1-1}),\ldots,(\eta_{i_\nu}-\eta_{i_\nu-1})$ dimensional vectors. The linear independence of $\{X_{i_1}^1\underline{\alpha}_1,\ldots,X_{i_\nu}^1\alpha_\nu\}$ then implies that the vectors

$\{\tilde{\underline{u}}_{1,i_1},\ldots,\tilde{\underline{u}}_{\nu,i_\nu}\}$ of η_1 dimension, are linearly independent.

In the case where $i_1=i_2=\ldots=i_\nu$, the independence of $\{\tilde{\underline{u}}_{1,i_1},\ldots,\tilde{\underline{u}}_{\nu,i_\nu}\}$ implies that $c_{\ell,P}=0$ for all $\ell=1,2,\ldots,k$, $P=i_1,\ldots,i_\nu$ (see Corollary (4.8)). In the case where at least one strict inequality holds true, the result is proved for a simple case and the general case follows along similar lines. Thus consider the case where $i_1=3$, $i_2=4$ and let the vectors defining the two chains respectively be $\underline{\alpha}_1$ and $\underline{\alpha}_2$. Then, by taking $k=4(\max\{i_1,i_2\})$ we have

$$\underline{P}(3,\underline{\alpha}_1)=\begin{bmatrix}P^3_1\underline{\alpha}_1\\P^3_2\underline{\alpha}_1\\P^3_3\underline{\alpha}_1\\0\end{bmatrix}=\begin{bmatrix}\underline{u}_{3,1}\\\underline{u}_{3,2}\\\underline{u}_{3,3}\\0\end{bmatrix},\quad \underline{P}(4,\underline{\alpha}_2)=\begin{bmatrix}P^4_1\underline{\alpha}_2\\P^4_2\underline{\alpha}_2\\P^4_3\underline{\alpha}_3\\P^4_4\underline{\alpha}_4\end{bmatrix}=\begin{bmatrix}\underline{u}_{4,1}\\\underline{u}_{4,2}\\\underline{u}_{4,3}\\\underline{u}_{4,4}\end{bmatrix} \tag{4.60a}$$

By (4.59c) and (4.59b) we obtain the following set of conditions

$$c_{1,3}\underline{u}_{3,1}+c_{1,4}\underline{u}_{4,1}+\begin{bmatrix}\underline{0}\\c_{2,3}\underline{u}_{3,2}+c_{2,4}\underline{u}_{4,2}\end{bmatrix}+\begin{bmatrix}\underline{0}\\\underline{0}\\c_{3,3}\underline{u}_{3,3}+c_{3,4}\underline{u}_{4,3}\end{bmatrix}+\begin{bmatrix}\underline{0}\\\underline{0}\\\underline{0}\\c_{4,4}\underline{u}_{4,4}\end{bmatrix}=\underline{0} \tag{4.60b}$$

$$c_{1,3}\underline{u}_{3,2}+c_{1,4}\underline{u}_{4,2}+\begin{bmatrix}\underline{0}\\c_{2,3}\underline{u}_{3,3}+c_{2,4}\underline{u}_{4,3}\end{bmatrix}+\begin{bmatrix}\underline{0}\\\underline{0}\\c_{3,4}\underline{u}_{4,4}\end{bmatrix}=\underline{0} \tag{4.60c}$$

$$c_{1,3}\underline{u}_{3,3}+c_{1,4}\underline{u}_{4,3}+\begin{bmatrix}\underline{0}\\c_{2,4}\underline{u}_{4,4}\end{bmatrix}=\underline{0} \tag{4.60d}$$

$$c_{1,4}\underline{u}_{4,4}=\underline{0} \tag{4.60e}$$

The assumption that $X^1_3\underline{\alpha}_1$, $X^1_4\underline{\alpha}_4$ are independent implies that

$$\underline{u}_{3,3} \neq \mu \begin{bmatrix} 0 \\ \underline{u}_{4,4} \end{bmatrix} , \ \mu\neq 0 <=> \begin{bmatrix} \underline{u}^{3}_{3,3} \\ \underline{u}^{4}_{3,3} \end{bmatrix} \neq \mu \begin{bmatrix} 0 \\ \underline{u}_{4,4} \end{bmatrix} \tag{4.60f}$$

The vectors in equations (4.60b-e) are partitioned next in the natural way implied by the form of these equations. Clearly, since $\underline{u}_{4,4}\neq 0$ (4.60e) implies that $c_{1,4}=0$. By setting $c_{1,4}=0$ in (4.60b-d) and using the partitioned form of the vectors, (4.60d) yields

$$c_{1,3}\underline{u}^{3}_{3,3} = 0, \ c_{1,3}\underline{u}^{4}_{3,3}+c_{2,4}\underline{u}_{4,4} = 0 \tag{4.61a}$$

(i) $\underline{\underline{u}^{3}_{3,3}\neq 0}$: (4.61a) implies $c_{1,3}=0$ and $c_{2,4}\underline{u}_{4,4}=0$ from which $c_{2,4}=0$ (since $\underline{u}_{4,4}\neq 0$). By setting $c_{1,3}=c_{2,4}=c_{1,4}=0$ into (4.60c) and using the partitioned form we have

$$c_{2,3}\underline{u}^{3}_{3,3} = 0, \ c_{2,3}\underline{u}^{4}_{3,3}+c_{3,4}\underline{u}_{4,4} = 0 \tag{4.61b}$$

from which $c_{2,3}=c_{3,4}=0$. By substitution in (4.60b) it is readily shown that $c_{3,3}=c_{4,4}=0$ and this leads to a contradiction.

(ii) $\underline{\underline{u}^{3}_{3,3}=0 \text{ and } \underline{u}^{4}_{3,3}\neq\mu\underline{u}_{4,4}}$: By the second of (4.61a) and (4.60f) we have that $c_{1,3}=c_{2,4}=0$ (otherwise the first vectors are dependent). By setting $c_{1,4}=c_{1,3}=c_{2,4}=0$ into (4.60c) we have

$$c_{2,3}\underline{u}^{3}_{3,3} = 0, \ c_{2,3}\underline{u}^{4}_{3,3}+c_{3,4}\underline{u}_{4,4} = 0 \tag{4.61c}$$

From the second of the above conditions and the (4.60f) condition we have $c_{2,3}=c_{3,4}=0$. Finally, by substituting into the reduced (4.60b) conditions we have

$$c_{3,3}\underline{u}^{3}_{3,3} = 0, \ c_{3,3}\underline{u}^{4}_{3,3}+c_{4,4}\underline{u}_{4,4} = 0 \tag{4.61d}$$

by (4.60f) then we have that $c_{3,3}=c_{4,4}=0$ and this once more leads to a contradiction.

The steps and arguments used in the simple case considered above are general; along similar lines it may be shown that the independence of $\{\tilde{\underline{u}}_{1,i_1},\dots,\tilde{\underline{u}}_{\nu,i_\nu}\}$ implies that the only solution of (4.59b) is the $c_{\ell,p}=0$ for $\forall\ \ell=1,2,\dots,k$, $p=i_1,\dots,i_\nu$ and thus sufficience of the result is established. □

The results so far suggest a procedure for selecting linear independent vector chains of the $S(i,\underline{\alpha}_i)$ type.

<u>Definition (4.4)</u>: Let N_α^τ be a nested basis matrix of $\mathcal{N}_\alpha^\tau$, as in (4.49), where τ is the index of annihilation of (F,G) at $s=\alpha$ and let $T_1,\dots,T_i,\dots,T_\tau$ be the matrices defined by

$$T_1=[X_\tau^1,X_{\tau-1}^1,\dots,X_i^1,\dots,X_2^1,X_1^1],\dots,\ T_i=[X_\tau^1,X_{\tau-1}^1,\dots,X_{i-1}^1,X_i^1],\dots$$
$$\dots,\ T_{\tau-1}=[X_\tau^1,X_{\tau-1}^1],\ T_\tau=[X_\tau^1] \qquad (4.62)$$

and $\mathcal{T}_i=\text{sp}\{T_i\}$ for all $i=1,\dots,\tau$. A basis $\mathcal{B}_\alpha=\{\underline{x}_1^{i_1},\dots,\underline{x}_{\omega_1}^{i_1};\dots;\underline{x}_1^{i_j},\dots,\underline{x}_{\omega_j}^{i_j};\dots;\underline{x}_1^{i_\mu},\dots,\underline{x}_{\omega_\mu}^{i_\mu}\}$ for $\mathcal{T}_1$ may be defined in the following way:

$\mathcal{B}_\alpha(i_1)=\{\underline{x}_1^{i_1},\dots,\underline{x}_{\omega_1}^{i_1}\}$ is a basis for $\mathcal{T}_\tau$ and $i_1=\tau$. If i_2 is the maximal index for which $\mathcal{T}_\tau=\mathcal{T}_{\tau-1}=\dots=\mathcal{T}_{i_2+1}\subset\mathcal{T}_{i_2}$, then $\mathcal{B}_\alpha(i_2)=\{\underline{x}_1^{i_2},\dots,\underline{x}_{\omega_2}^{i_2}\}$ is a maximal set of independent column vectors in T_{i_2} which do not belong to $\mathcal{T}_{i_1}$. Similarly, let i_3 be the maximal index for which $\mathcal{T}_{i_2}=\mathcal{T}_{t_2+1}=\dots=\mathcal{T}_{i_3+1}\subset\mathcal{T}_{i_3}$; then let $\mathcal{B}_\alpha(i_3)=\{\underline{x}_1^{i_3},\dots,\underline{x}_{\omega_3}^{i_3}\}$ be the maximal set of independent column vectors in T_{i_3}, which are not in $\mathcal{T}_{i_2}$. The procedure eventually terminates for some index i_μ for which we have that $\dots=\mathcal{T}_{i_\mu-2}=\mathcal{T}_{i_\mu+1}\subset\mathcal{T}_{i_\mu}=\mathcal{T}_{i_\mu-1}=\dots=\mathcal{T}_1$; then $\mathcal{B}_\alpha(i_\mu)=\{\underline{x}_1^{i_\mu},\dots,\underline{x}_{\omega_\mu}^{i_\mu}\}$ is a maximal set of independent column vectors in T_{i_μ} which are not in $\mathcal{T}_{i_\mu-1}$. The basis $\mathcal{B}_\alpha$ for $\mathcal{T}_1$

$$\begin{aligned}\mathcal{B}_\alpha &= \{\mathcal{B}_\alpha(i_1);\dots;\mathcal{B}_\alpha(i_j);\dots;\mathcal{B}_\alpha(i_\mu)\}\\ &= \{\underline{x}_1^{i_1},\dots,\underline{x}_{\omega_1}^{i_1};\dots;\underline{x}_1^{i_j},\dots,\underline{x}_{\omega_j}^{i_j};\dots;\underline{x}_1^{i_\mu},\dots,\underline{x}_{\omega_\mu}^{i_\mu}\}\end{aligned} \qquad (4.63)$$

which has been constructed with the procedure described above, will be called a <u>normal basis of generators</u> of (F,G) at $s=\alpha$. The ordered set of indices $L=\{(i_1,\omega_1),\ldots,(i_j,\omega_j),\ldots,(i_\mu,\omega_\mu)\colon i_1=\tau>i_2>\ldots>i_j>\ldots>i_\mu\}$ will be referred to as the list of B_α and every vector in $B_\alpha(i_j)$ will be called an i_j-th order generator of (F,G) at $s=\alpha$. Every i_j-th order generator $\underline{x}_k^{i_j}$ in $B_\alpha(i_j)$ defines a column in N_α^τ, and by partitioning this column vector in the natural way implied by the partitioning of N_α^τ we obtain an i_j-th length independent chain of vectors $S_\alpha(i_j,\underline{x}_k^{i_j})=\{\underline{x}_{k,1}^{i_j},\ldots,\underline{x}_{k,i_j}^{i_j}\colon \underline{x}_{k,1}^{i_j}=\underline{x}_k^{i}\}$ for which

$$G\underline{x}_{k,1}^{i_j} = \alpha F\underline{x}_{k,1}^{i_j},\ G\underline{x}_{k,2}^{i_j} = \alpha F\underline{x}_{k,2}^{i_j}+F\underline{x}_{k,1}^{i_j},\ldots$$

$$\ldots,G\underline{x}_{k,i_j}^{i_j} = \alpha F\underline{x}_{k,i_j}^{i_j}+F\underline{x}_{k,i_j-1}^{i_j} \tag{4.64}$$

$S_\alpha(i_j,\underline{x}_k^{i_j})$ will be called an i_j-th order prime chain of (F,G) at $s=\alpha$ and the space defined by $S_{i_j}^k=\mathrm{sp}\{\underline{x}_{k,1}^{i_j},\ldots,\underline{x}_{k,i_j}^{i_j}\}$ is i_j-dimensional and will be called an i_j-th order prime subspace of (F,G) at $s=\alpha$. The set of all prime chains of all possible orders will be denoted by $\Sigma_\alpha(F,G)=\{S_\alpha(i_1,\underline{x}_1^{i_1}),\ldots,S_\alpha(i_1,\underline{x}_{\omega_1}^{i_1});\ldots;S_\alpha(i_\mu,\underline{x}_1^{i_\mu}),\ldots,S_\alpha(i_\mu,\underline{x}_{\omega_\mu}^{i_\mu})\}$ and the set of all prime subspaces of all possible order by $S_\alpha(F,G)=\{S_{i_1}^1,\ldots,S_{i_1}^{\omega_1};\ldots;S_{i_\mu}^1,\ldots,S_{i_\mu}^{\omega_\mu}\}$; $\Sigma_\alpha(F,G)$ and $S_\alpha(F,G)$ will be called respectively a complete prime set of chains and a complete prime set of subspaces of (F,G) at $s=\alpha$.

An obvious property of every i_j-th order prime chain, following from the fact that N_α^τ is a basis matrix of the maximal dimension N_α^i type space is given next.

Remark (4.11): Every i_j-th order prime chain $S_\alpha(i_j,\underline{x}_k^{i_j})$ is of maximal length i_j. The set $\Sigma_\alpha(F,G)$ is made up from maximal chains with possible lengths $i_1,i_2,\ldots,i_\mu$.

With these definitions in mind, we may give the following result summarising the properties of the $S_\alpha(F,G)$ and $\Sigma_\alpha(F,G)$ sets.

Theorem (4.5): Let N_{α}^{τ} be a nested basis of $\mathcal{N}_{\alpha}^{\tau}$, τ is the index of annihilation of (F,G) at $s=\alpha$, $\mathcal{B}_{\alpha}$ a corresponding normal basis of generators, with an associated list $\mathcal{L}=\{(i_1,\omega_1),\ldots,(i_{\mu},\omega_{\mu}): i_1=\tau>i_2>\ldots>i_{\mu}\}$, and let $\mathcal{S}_{\alpha}(F,G)$ be the corresponding complete prime set of subspaces of (F,G) at $s=\alpha$. Then, we have the following properties:

(i) The list $\mathcal{L}$ is an invariant of all nested basis matrices of $\mathcal{N}_{\alpha}^{\tau}$.

(ii) If $\mathcal{I}=\{(d_1,\sigma_1),\ldots,(d_{\rho},\sigma_{\rho}): d_1<\ldots<d_{\rho}\}$ is the index set of (F,G) at $s=\alpha$, then $\mu=\rho$ and the sets $\mathcal{I},\mathcal{L}$ are related as follows:

$$(d_1,\sigma_1)=(i_{\rho},\omega_{\rho}),(d_2,\sigma_2)=(i_{\rho-1},\omega_{\rho-1}),\ldots,(d_{\rho},\sigma_{\rho})=(i_1,\omega_1) \tag{4.65}$$

where by $(d,\sigma)\overset{\Delta}{=}(i,\omega)$ we mean that $d=i$ and $\sigma=\omega$.

(iii) Any complete prime set of subspaces $\mathcal{S}_{\alpha}(F,G)$, or any complete prime set of chains $\Sigma_{\alpha}(F,G)$, is linearly independent.

(iv) If $\mathcal{M}_{\alpha}^{*}$ is the maximal generalised nullspace of (F,G) at $s=\alpha$, then $\mathcal{M}_{\alpha}^{*}$ may be expressed in terms of the i_j-th order prime subspaces $\mathcal{S}_{i_j}^{k}$ of any set $\mathcal{S}_{\alpha}(F,G)$ as

$$\mathcal{M}_{\alpha}^{*}=\mathcal{S}_{i_1}^{1}\oplus\ldots\oplus\mathcal{S}_{i_1}^{\omega_1}\oplus\ldots\oplus\mathcal{S}_{i_{\mu}}^{1}\oplus\ldots\oplus\mathcal{S}_{i_{\mu}}^{\omega_{\mu}} \tag{4.66}$$

Proof

(i) Let $N_{\alpha}^{\tau},\tilde{N}_{\alpha}^{\tau}$ be two nested basis matrices of $\mathcal{N}_{\alpha}^{\tau}$. By Proposition (4.7) $\tilde{N}_{\alpha}^{\tau}=N_{\alpha}^{\tau}P_{\tau})$ and thus the blocks on the main nonzero diagonal are related by $\tilde{X}_{i}^{1}=X_{i}^{1}P_{i}^{i}$, $i=1,2,\ldots,\tau$, where P_{i}^{i} are square nonsingular matrices. Then,

$$\tilde{T}_{i}=[\tilde{X}_{\tau}^{1},\tilde{X}_{\tau-1}^{1},\ldots,\tilde{X}_{i}^{1}]=[X_{\tau}^{1},X_{\tau-1}^{1},\ldots,X_{i}^{1}]\mathrm{diag}\{P_{\tau}^{\tau},P_{\tau-1}^{\tau-1},\ldots,P_{i}^{i}\}$$

$$i=1,2,\ldots,\tau \tag{4.67a}$$

By (4.67a) and the selection procedure given in Definition (4.4), it is obvious that the lists of $\mathcal{B}_{\alpha}$ and $\tilde{\mathcal{B}}_{\alpha}$ are the same.

(ii) Since the list is an invariant of any nested basis of $\mathcal{N}_{\alpha}^{\tau}$, we may use a particular basis for the exploration of the links of $\mathcal{L}$ and $\mathcal{I}$. Choose as a nested basis the $\hat{N}_{\alpha}^{\tau}=\text{diag}\{Q,\ldots,Q\}E_{\tau}$. Then we have that

$$\hat{T}_1 = Q[E_{\tau}^1, E_{\tau-1}^1, \ldots, E_i^1, \ldots, E_1^1] = QT_e \tag{4.67b}$$

where $E_i^1 \overset{\Delta}{=} E^1(\phi_i)$. Clearly, the lists computed on $\hat{T}_1$ and T_e are the same; thus, we may use T_e for the computation of the list of any T_1 matrix. From the definition of the E_i^1 matrices we have that $\text{rank}\{E_i^1\}=\text{rank}\{E_i^i\}=\eta_i-\eta_{i-1}$; thus for the (i_1,ω_1) pair we have that $i_1=\tau=d_{\rho}$ and $\omega_1=\eta_{\tau}-\eta_{\tau-1}=\sigma_{\rho}$. If i_2 is the maximal index for which $\text{sp}\{E_{\tau}^1\}=\text{sp}\{E_{\tau-1}^1\}=\ldots=\text{sp}\{E_{i_2+1}^1\}\subset\text{sp}\{E_{i_2}^1\}$, then we have that

$$E_{\tau}^1 = E_{\tau-1}^1 = \ldots = E_{i_2+1}^1 \subset^c E_{i_2}^1 \tag{4.67c}$$

where $\subset^c$ implies that the columns of $E_{i_2+1}^1$ are strictly contained in the columns of $E_{i_2}^1$, and that

$$\eta_{\tau}-\eta_{\tau-1}=\eta_{\tau-1}-\eta_{\tau-2}=\ldots=\eta_{i_2+1}-\eta_{i_2}<\eta_{i_2}-\eta_{i_2-1} \tag{4.67d}$$

By Proposition (4.5) and conditions (4.67d), it follows that i_2 is singular point of the piecewise arithmetic progression sequence $\{\eta_o,\eta_1,\ldots,\eta_{i_2},\eta_{i_2+1},\ldots,\eta_{\tau}\}$ and thus $i_2=d_{\rho-1}$. By condition (4.67c) the maximal number of independent vectors in T_{i_2}, which are not in T_{i_2+1} is equal to $\omega_2=(\eta_{i_2}-\eta_{i_2-1})-(\eta_{i_2+1}-\eta_{i_2})=2\eta_{i_2}-\eta_{i_2-1}-\eta_{i_2+1}=\delta_{i_2}$, where δ_{i_2} is the gap of the sequence at i_2. By Proposition (4.5), we have that $\sigma_{\rho-1}=\delta_{d_{\rho-1}}=\delta_{i_2}=\omega_2$. By repeating the arguments used in the above step, it follows that $\mu=\rho$ and that

$$(d_{\rho-2},\sigma_{\rho-2}) = (i_3,\omega_3),\ldots,(d_1,\omega_1) = (i_{\rho},\omega_{\rho})$$

(iii) A complete prime set of chains $\Sigma_{\alpha}(F,G)$ is generated by the vectors of $\mathcal{B}_{\alpha}$. Since the vectors in $\mathcal{B}_{\alpha}$ are linearly independent, then by Proposition (4.10) the chains in $\Sigma_{\alpha}(F,G)$ are independent and the set of subspaces $S_{\alpha}(F,G)$ is also independent.

(iv) Since the set $S_{\alpha}(F,G)$ is linearly independent, in order to prove the direct sum decomposition (4.66) we have to show that

$$M_{\alpha}^{*} = \sum_{j=1}^{\omega_1} S_{i_1}^{j} + \ldots + \sum_{j=1}^{\omega_\mu} S_{i_\mu}^{j}$$

By Proposition (4.9) it follows that for $\forall\ k \in \underset{\sim}{\mu}$ and $j \in \underset{\sim}{\omega_k}$ we have that $S_{i_k}^{j} \subseteq M_{\alpha}^{*}$; given that $S_{i_k}^{j}$ are linearly independent for all $k \in \underset{\sim}{\mu}$ and $j \in \underset{\sim}{\omega_k}$ we have that the subspace

$$W = \{ \bigoplus_{j=1}^{\omega_1} S_{i_1}^{j} \} \oplus \ldots \oplus \{ \bigoplus_{j=1}^{\omega_\mu} S_{i_\mu}^{j} \} \subseteq M_{\alpha}^{*} \tag{4.67e}$$

and that $\dim W = \sum_{k=1}^{\mu} \sum_{j=1}^{\omega_k} S_{i_k}^{j} = \sum_{j=1}^{\mu} i_j \omega_j$. By part (ii) of the theorem and the last expression for $\dim W$, we also have that $\dim W = \sum_{i=1}^{\rho} d_i \sigma_i$, where (d_i, σ_i) are the element of I. However, by Corollary (4.4), $\eta_{\tau} = \sum_{i=1}^{\rho} d_i \sigma_i = \dim W$, and since $\eta_{\tau} = \dim M_{\alpha}^{*}$, we have $\dim M_{\alpha}^{*} = \dim W$. The conditions $\dim M_{\alpha}^{*} = \dim W$ and $W \subseteq M_{\alpha}^{*}$ clearly imply that $W = M_{\alpha}^{*}$. □

The above result suggests an alternative procedure for the computation of the index set I of (F,G) at $s=\alpha$; a procedure for finding a set of linearly independent maximal prime chains of vectors characterising the set of e.d. at $s=\alpha$, is also suggested by Theorem (4.5). We may summarise this procedure as follows:

Nested basis matrix approach: Let τ be the index of annihilation of (F,G) at $s=\alpha$ and let N_{α}^{τ} be a nested basis matrix for N_{α}^{τ}. Following the steps suggested by Definition (4.4) we find:

(i) A normal basis for generators $B_{\alpha} = \{B_{\alpha}(i_1);\ldots;B_{\alpha}(i_j);\ldots;B_{\alpha}(i_\mu)\}$, the associated list of B_{α}, $L = \{(i_1,\omega_1),\ldots,(i_j,\omega_j),\ldots,(i_\mu,\omega_\mu) : i_1 = \tau > i_2 > \ldots \ldots > i_j > \ldots > i_\mu\}$, and the corresponding complete prime set of chains $\Sigma_{\alpha}(F,G) = \{S_{\alpha}(i_1, \underline{x}_1^{i_1}),\ldots,S_{\alpha}(i_1, \underline{x}_{\omega_1}^{i_1});\ldots;S_{\alpha}(i_\mu, \underline{x}_1^{i_\mu}),\ldots,S_{\alpha}(i_\mu, \underline{x}_{\omega_\mu}^{i_\mu})\}$, generated by B_{α}.

(ii) By reordering the elements of L in ascending order of the i_j and renaming them, we obtain I as follows: set $\mu=\rho$ and define

$$(d_1,\sigma_1)=(i_\rho,\omega_\rho),(d_2,\sigma_2)=(i_{\rho-1},\omega_{\rho-1}),\dots,(d_\rho,\sigma_\rho)=(i_1=\tau,\omega_1)$$

(iii) By reordering the elements of $\Sigma_\alpha(F,G)$ in ascending order of the i_j and by using I for its parametrisation we obtain the set

$$\tilde{\Sigma}_\alpha(F,G)=\{S_\alpha(d_1,\underline{x}_1^{d_1}),\dots,S_\alpha(d_1,\underline{x}_{\sigma_1}^{d_1});\dots;S_\alpha(d_\rho,\underline{x}_1^{d_\rho}),\dots,S_\alpha(d_\rho,\underline{x}_{\sigma_\rho}^{d_\rho})\}$$

$\tilde{\Sigma}_\alpha(F,G)$ will be referred to as <u>normal complete prime set of chains</u> of (F,G) at $s=\alpha$. Every chain $S_\alpha(d_i,\underline{x}_j^{d_i})$, $j\in\underline{\sigma}_i$, is a maximal chain of linearly independent vectors characterising an e.d. $(s-\alpha)^{d_i}$; the chain in $\tilde{\Sigma}_\alpha(F,G)$ are linearly independent and the associated subspaces $S_{d_i}^j$ provide a direct sum decomposition for M_α^* as in (4.66). □

This third approach, based on a nested basis matrix, has the advantage that apart from the computation of I, also yields the set $\tilde{\Sigma}_\alpha(F,G)$ of vector chains characterising the set of e.d. of (F,G) at $s=\alpha$; thus, on one hand provides the means for the computation of the Weierstrass canonical form, and on the other hand indicates the procedure for the derivation of a transformation pair (R,Q) that reduces $(F,G)\xrightarrow{(R,Q)}(F_w,G_w)$. We conclude this section by giving a result that suggests how we can construct a nested basis matrix of N_α^τ from any basis matrix of N_α^τ.

<u>Proposition (4.11)</u>: The column echelon form basis matrix, H_α^c, of N_α^τ which has its column ordered from right to left, is a canonical nested basis matrix of N_α^τ.

<u>Proof</u>

Let $\hat{P}_\alpha^\tau,N_\alpha^\tau\in\mathbb{C}^{\tau n\times\eta_\tau}$ be two right annihilators of $P_\alpha^\tau(F,G)$, where τ is the index of annihilation of (F,G) at $s=\alpha$, and N_α^τ is a nested basis matrix. Then there exists $T\in\mathbb{C}^{\eta_\tau\times\eta_\tau}$, $|T|\neq0$, such that $N_\alpha^\tau=\hat{P}_\alpha^\tau T$, and thus $N_\alpha^\tau,\hat{P}_\alpha^\tau$ are right (column) equivalent and the equivalence class of all right annihilators of $\mathcal{N}_{\alpha}^{\tau}$ is characterised by a unique column Hermite form (column echelon form) [Mar & Min - 1]. To construct the echelon form we may start from N_{α}^{τ}. The standard procedure for the construction of the column echelon form may be applied on the column blocks of N_{α}^{τ}, with the only difference that we start from the last block, the τ-th order and we go backward to the 1-st order block. It is readily verified that every column operation used does not affect the structure of the nested basis, and thus the reduction of N_{α}^{τ} to H_{α}^{c} is achieved by transformations of the type (4.38a,b). □

Remark (4.12): A nested basis matrix N_{α}^{τ} for $\mathcal{N}_{\alpha}^{\tau}$ may be constructed from any basis matrix $\hat{P}_{\alpha}^{\tau}$ of $\mathcal{N}_{\alpha}^{\tau}$ by the type of elementary column operations used for the reduction of $\hat{P}_{\alpha}^{\tau}$ to its column echelon form, which has its columns ordered from right to left.

The analysis so far has been restricted to the case of a single frequency $s=\alpha$; clearly, the results are the same for the case of $s=\infty$, with the only difference that the matrices $P_{\infty}^{i}(F,G)$ are now considered instead of the $P_{\alpha}^{i}(F,G)$ matrices. We close this chapter by discussing some properties of the set of subspaces $\mathcal{M}_{\alpha}^{*}$, where $\alpha\in\mathbb{C}\cup\{\infty\}$ and $\alpha F-G$ singular. The results presented next provide alternative techniques for the derivation of the Weierstrass canonical form of a regular pencil.

4.5 On the derivation of Weierstrass canonical form of a regular pencil

The key notion for the derivation of the Weierstrass canonical form is the index set $\mathcal{I}_{\alpha}$, or equivalently the Segre characteristic $\mathcal{S}_{\alpha}$ of (F,G) at $s=\alpha$. The computation of the Weierstrass form may be achieved, without finding the pair of transformations (R,Q) which reduce (F,G) to (F_w,G_w); however, in a number of applications it is important to know also a pair (R,Q), apart from the canonical form itself. In the following, those two problems are considered.

<u>Segre′ characteristic based approach</u>: For the n×n regular pencil sF-G the Weierstrass form sF_w-G_w may be computed <u>without</u> finding a pair (R,Q) for which $R(sF-G)Q=sF_w-G_w$. The suggested procedure is as follows:

(i) Find $sF-G = c_m s^m + c_{m-1}s^{m-1}+\ldots+c_0$, where $m \le n$, and find the roots of sF-G , with multiplicities included, say the set $\{(\alpha_1,\pi_1),\ldots,(\alpha_\mu,\pi_\mu)\}$, where α_i is a root and π_i is the corresponding algebraic multiplicity of the root. If m<n, then sF-G loses rank at $s=\infty$, and $|sF-G|$ has a root at $s=\infty$ with algebraic multiplicity $\pi_\infty=n-m$. The set $\bar{R}=\{(\infty,\pi_\infty),(\alpha_1,\pi_1),\ldots,$ $\ldots,(\alpha_\mu,\pi_\mu)\}$ will be called the <u>root set</u> of (F,G), whereas the set of distinct roots $R=\{\infty,\alpha_1,\ldots,\alpha_\mu\}$ will be called the <u>root range</u> of (F,G).

(ii) For every $\beta\in R$ compute the matrices $P^i_\beta(F,G)$ and dimensions of $N_r\{P^i_\beta(F,G)\}$, i=1,2,..., say η^β_i, i=1,2,... From the numbers $\{\eta^\beta_1,\eta^\beta_2,\ldots,\eta^\beta_i,\ldots\}$ find the Segre characteristic by using the Ferrer's diagram, or the Piecewise Arithmetic Progression Sequence diagram. Note that since the set $\{\alpha_1,\ldots,\alpha_\mu\}$ is symmetric (the complex numbers are in complex conjugate pairs), the computations are carried out for ∞ , the real elements of R and for only one number of every complex conjugate pair $(\alpha_i, \bar{\alpha}_i)$.

(iii) The set $S_e(F,G)=\{S_\infty,S_{\alpha_1},\ldots,S_{\alpha_\mu}\}$, where S_β is the Segre′ characteristic of (F,G) at $s=\beta$, $\beta\in R$, is called the <u>Segre′ characteristic</u> of (F,G) and defined the Weierstrass form of (F,G). Note if τ_β is the index of annihilation of (F,G) at $s=\beta$, then $\eta_{\tau_\beta}=\pi_\beta$.

<u>Remark (4.13)</u>: For the derivation of the Weierstrass form with the procedure suggested above, the set R is needed for the computations, i.e. the set of distinct complex numbers in $\mathbb{C}\cup\{\infty\}$ for which sF-G loses rank, and not $\bar{R}$. If $\bar{R}$ is available, then it may be used for the computation of τ_β, since τ_β will be the index for which $\eta_{\tau_\beta}=\pi_\beta$.

The procedure discussed so far for the computation of (F_w,G_w) does not indicate how a pair (R,Q), such that $R(sF-G)Q=sF_w-G_w$ may be computed. To define such a pair of transformations we have to use the properties of the nested basis matrices discussed in section (4.4). Before we proceed to the discussion of an alternative procedure for the derivation of the Weierstrass form we give the following useful result.

Theorem (4.6): Let $\mathcal{R}=\{\infty,\alpha_1,\dots,\alpha_\mu\}$ be the root range of (F,G) and let $\mathcal{M}_\beta^*$ be the maximal generalised nullspace of (F,G) at $s=\beta$, $\beta\epsilon\mathcal{R}$. The following properties hold true:

(i) The set of subspaces of $\mathbb{C}^n$ $\{\mathcal{M}_\infty^*,\mathcal{M}_{\alpha_1}^*,\dots,\mathcal{M}_{\alpha_\mu}^*\}$ are linearly independent.

(ii) $\mathbb{C}^n=\mathcal{M}_\infty^*\oplus\mathcal{M}_{\alpha_1}^*\oplus\dots\oplus\mathcal{M}_{\alpha_\mu}^*$.

Proof

Let (R,Q) be a pair of transformations which reduce (F,G) to its Weierstrass form (F_w,G_w), i.e.

$$R(sF-G)Q=sF_w-G_w=\text{diag}\{D(\infty);D(\alpha_1);\dots;D(\alpha_\mu)\} \tag{4.68a}$$

where

$$D(\infty)=\text{diag}\{I_{q_1}-sH_{q_1};\dots;I_{q_p}-sH_{q_p}\}$$

$$D(\alpha_i(=\text{diag}\{sI_{d_{i,1}}-J_{d_{i,1}}(\alpha_i);\dots;sI_{d_i,\nu_i}-J_{d_i,\nu_i}(\alpha_i)\},\quad i\epsilon\underset{\sim}{\mu} \tag{6.68b}$$

If we partition Q according to the partitioning of sF_w-G_w in (4.68a), i.e. $Q=[Q_\infty,Q_1,\dots,Q_\mu]$, then clearly the set of subspaces defined by $\tau_\infty=\text{col.-sp}\{Q_\infty\},\tau_{\alpha_i}=\text{col.-sp}\{Q_i\}$, $i\epsilon\underset{\sim}{\mu}$ are linearly independent, since $Q_\infty,Q_1,\dots,Q_\mu$ are column blocks of the full rank matrix Q. We shall show that $\tau_\infty=\mathcal{M}_\infty^*,\tau_{\alpha_i}=\mathcal{M}_{\alpha_i}^*$, $i\epsilon\underset{\sim}{\mu}$ and thus part (i) and (ii) of the result will be evident. The result will be proved for a general α, whereas the proof for $\alpha=\infty$ is similar.

Let $\mathcal{I}_\alpha=\{(d_1,\sigma_1),\dots,(d_\rho,\sigma_\rho),d_1<\dots<d_\rho\}$ be the index set of (F,G) at $s=\alpha$ and let Q_α be the column block of Q that corresponds to $D(\alpha)$. The block

diagonal structure of $D(\alpha)$ (as in (4.68b)) implies that Q_α may be partitioned as

$$Q_\alpha = \left[Q_\alpha^{1,d_1};\ldots;Q_\alpha^{\sigma_1,d_1};\ldots;Q_\alpha^{1,d_\rho};\ldots;Q_\alpha^{\sigma_\rho,d_\rho} \right] \tag{4.68c}$$

If we now write every column block Q_α^{j,d_i}, $j\epsilon\underset{\sim}{\sigma}_i$, $i\epsilon\underset{\sim}{\rho}$ into terms of its columns, i.e. $Q_\alpha^{j,d_i}=[\underline{x}_1^{j,i},\ldots,\underline{x}_{d_i}^{j,i}]$, then by the proof of Proposition (4.1) we have that

$$(G-\alpha F)\underline{x}_k^{j,i} = F\underline{x}_{k-1}^{j,i},\ k\epsilon\underset{\sim}{d}_i,\ \underline{x}_o^{j,i}=0,\ j\epsilon\underset{\sim}{\sigma}_i,\ i\epsilon\underset{\sim}{\rho} \tag{4.68d}$$

From the set of linearly independent vectors $\{\underline{x}_k^{j,i},k\epsilon\underset{\sim}{d}_i,j\epsilon\underset{\sim}{\sigma}_i,i\epsilon\underset{\sim}{\rho}\}$ we can construct a matrix Y_α of dimension $n\tau_\alpha \times \sum_{i=1}^{\rho}\sigma_i$ and with the structure of a nested basis matrix (eqn(4.37)), where the various blocks are formed in the following way:

(i) $X_1^1=[\underline{x}_1^{1,1},\ldots,\underline{x}_1^{\sigma_1,1};\ldots;\underline{x}_1^{1,\rho},\ldots,\underline{x}_1^{\sigma_\rho,\rho}]$, i.e. all vectors $\underline{x}_1^{j,i}$ arranged in increasing order of d_i.

(ii) Find all vector chains of length greater or equal to 2, and consider the set of all ordered pairs $(\underline{x}_1^{j,i},\underline{x}_2^{j,i})$ from these chains. The matrices X_2^2, X_2^1 are then constructed as

$$\begin{bmatrix} X_2^1 \\ \hline X_2^2 \end{bmatrix} = \left[\ldots, \begin{bmatrix} \underline{x}_1^{j,i} \\ \hline \underline{x}_2^{j,i} \end{bmatrix}, \ldots \right] \quad \text{, where } d_i \geq 2$$

and the columns are arranged in increasing order of d_i.

(iii) The set of matrices $X_k^k, X_k^{k-1}, \ldots X_k^1$, $k=1,\ldots,\rho$ are constructed as follows: Find all vector chains of length greater or equal to k, and consider the set of all ordered k-tuples $(\underline{x}_1^{j,i},\underline{x}_2^{j,i},\ldots,\underline{x}_{k-1}^{j,i},\underline{x}_k^{j,i})$ from these chains. Then construct the matrix

$$\begin{bmatrix} X_k^1 \\ \hdashline X_k^2 \\ \hdashline \vdots \\ \hdashline X_k^{k-1} \\ \hdashline X_k^k \end{bmatrix} = \left[\ldots, \begin{bmatrix} \underline{x}_1^{j,i} \\ \hdashline \underline{x}_2^{j,i} \\ \hdashline \vdots \\ \hdashline \underline{x}_{k-1}^{j,i} \\ \hdashline \underline{x}_k^{j,i} \end{bmatrix}, \ldots \right], \text{ where } d_i \geq k$$

and the columns are arranged in increasing order of d_i.

The matrix Y_α constructed with the above procedure has the following properties: it has dimensions $n\tau_\alpha \times \sum_{i=1}^{\rho} \sigma_i$, $\text{rank}(Y_\alpha) = \sum_{i=1}^{\rho} \sigma_i$ (since the set $\{\underline{x}_k^{j,i}, k \in \underset{\sim}{d}_i, j \in \underset{\sim}{\sigma}_i, i \in \underset{\sim}{\rho}\}$ is linearly independent) and every column of Y_α is in $\mathcal{N}_r\{P_\alpha^{\tau_\alpha}(F,G)\}$ (by conditions (4.68d)); thus, Y_α is a nested basis matrix of (F,G) at $s=\alpha$. By construction the matrices $Y_\alpha^1 = [X_1^1, X_2^2, \ldots, X_\rho^\rho]$, $\rho = \tau_\alpha$, and Q_α have the same set of columns, arranged in different order; thus $\text{col-sp}\{Y_\alpha^1\} = \mathcal{M}_\alpha^* = \text{col-sp}\{Q_\alpha\} = \tau_\alpha$ and the result is established. □

The problem of finding a pair $(R,Q), R, Q \in \mathbb{C}^{n \times n}, |R|, |Q| \neq 0$, such that $(F,-G) \xrightarrow{(R,Q)} (RFQ,-RGQ) = (F_w -G_w)$ is considered next. This problem may be expressed in matrix form by the equation

$$R[F,-G] \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = [F_w, -G_w] \tag{4.69}$$

or equivalently by the matrix equation

$$[FQ, -GQ] = \hat{R}[F_w, -G_w] \tag{4.70}$$

where $\hat{R} = R^{-1}$, $(F,-G)$ the given pair, $(F_w, -G_w)$ its Weierstrass form and $(Q,\hat{R})$ are the unknown n×n matrices which must be of full rank. The solvability of this equation is characterised by the following result.

<u>Corollary (4.9)</u>: Necessary and sufficient condition for the existence of a solution pair $(\hat{R},Q)$ of eqn(4.70), where $\hat{R}, Q$ are n×n full rank matrices, is that Q is expressed as

$$Q = [Q_\infty, Q_{\alpha_1}, \ldots, Q_\beta, \ldots, Q_{\alpha_\mu}] \tag{4.71}$$

where the columns of Q_β form a complete prime set of chains of (F,G) at $s=\beta$, for every $\beta \epsilon \mathcal{R}$. If Q is chosen as above and if $\begin{bmatrix} \tilde{F}_w \\ -\tilde{G}_w \end{bmatrix}$ is a right inverse $[F_w, -G_w]$, then $\hat{R}=R^{-1}$ is given by

$$\hat{R} = FQ\tilde{F}_w + GQ\tilde{G}_w \tag{4.72}$$

Proof

(Necessity): If (R,Q) reduces (F,G) to (F_w, G_w), then by the Proof of Theorem (4.6), the columns of the Q_β column blocks of Q, for every $\beta \epsilon \mathcal{R}$, define a nested basis matrix Y_β for $\mathcal{N}_r\{P_\beta^{\tau_\beta}(F,G)\}$. If we now start from Y_β, it is clear from the conditions (4.68c) and (4.68d) that the columns of Q_β form a complete prime set of chains of (F,G) at $s=\beta$ and this proves the necessity.

(Sufficiency): Assume that Q is selected as in (4.71), where for every $\beta \epsilon \mathcal{R}$ the columns of Q_β form a complete prime set of chains of (F,G) at $s=\beta$. Because of this property we have that the following conditions hold true:

(i) If $\beta \epsilon \{\alpha_1, \ldots, \alpha_\mu\}$, then

$$GQ_\beta = FQ_\beta J_{p_\beta}(\beta) = FQ_\beta \operatorname{diag}\{J_{\delta_1}(\beta), \ldots, J_{\delta_\theta}(\beta)\} \tag{4.73a}$$

where $\{\delta_1, \ldots, \delta_\theta\}$ are the degrees of the e.d. at $s=\beta$, $J_{\delta_i}(\beta)$ denote standard Jordan blocks at β of dimensions $\delta_i \times \delta_i$ and $p_\beta = \Sigma\delta_i$.

(ii) If $\beta=\infty$, then

$$FQ_\infty = GQ_\infty J_{q_\infty}(\infty) = GQ_\infty \operatorname{diag}\{J_{q_1}(0), \ldots, J_{q_\omega}(0)\} \tag{4.73b}$$

where $\{q_1, \ldots, q_\omega\}$ are the degrees of the infinite e.d., $J_{q_j}(0)$ denote standard Jordan blocks at 0 of dimensions $q_j \times q_j$ and $q_\infty = \Sigma q_j$.

Before we begin the study of solvability of eqn(4.70) for a given matrix Q satisfying conditions (4.73a) and (4.73b) we note that both $sF-G$, sF_w-G_w

are regular and thus rank $[FQ,-GQ]=n=\text{rank}[F_w,-G_w]$. Thus, if a solution $\hat{R}$ exists, then it is a nonsingular matrix. Since $T_w=[F_w,-G_w]\in\mathbb{C}^{n\times 2n}$ and has rank n, a right inverse $T_w^{\dagger}\in\mathbb{C}^{2n\times n}$ exists, that is $T_wT_w^{\dagger}=I_n$. If $T_w^{\perp}$ is a right annihilator of T_w, $T_w^{\perp}\in\mathbb{C}^{2n\times n}$, rank$\{T_w^{\perp}\}$=n and $T_wT_w^{\perp}=0$, then it is known that $U=[T_w^{\dagger},T_w^{\perp}]\in\mathbb{C}^{2n\times 2n}$ and has full rank (2n). By multiplying both sides of (4.70) on the right by U we reduce (4.70) to the following equivalent set of conditions

$$[FQ,-GQ]T_w^{\perp} = 0 \tag{4.74a}$$

$$\hat{R} = [FQ,-GQ]T_w^{\dagger} \tag{4.74b}$$

Clearly (4.74a) is necessary and sufficient for the solvability of (4.70), since if (4.74a) is satisfied, then $\hat{R}$ is defined by (4.74b).

In order to check (4.74a) we must compute a right annihilator of T_w. From the special structure of T_w, indicated below,

$$T_w = [F_w,-G_w] = \left[\begin{array}{cccc:cccc} J_{q_\infty}(\infty) & & \bigcirc & & -I_{q_\infty} & & & \bigcirc \\ & I_{p_{\alpha_1}} & & & & -J_{p_{\alpha_1}}(\alpha_1) & & \\ \bigcirc & & \ddots & & & & \ddots & \\ & & & I_{p_{\alpha_\mu}} & \bigcirc & & & -J_{p_{\alpha_\mu}}(\alpha_\mu) \end{array}\right] \tag{4.75a}$$

we can construct a matrix A as

$$A = \left[\begin{array}{c} F_w' \\ \hdashline -G_w' \end{array}\right] = \left[\begin{array}{cccc} I_{q_\infty} & & & \bigcirc \\ & J_{p_{\alpha_1}}(\alpha_1) & & \\ \bigcirc & & \ddots & \\ & & & J_{p_{\alpha_\mu}}(\alpha_\mu) \\ \hdashline J_{q_\infty}(\infty) & & & \bigcirc \\ & I_{p_{\alpha_1}} & & \\ \bigcirc & & \ddots & \\ & & & I_{p_{\alpha_\mu}} \end{array}\right] \tag{4.75b}$$

Clearly $A\in\mathbb{C}^{2n\times n}$, rank$\{A\}$=n and $T_wA=0$; thus, A is a right annihilator of T_w and in (4.74a) we may set $T_w^{\perp}=A$. With this choice of $T_w^{\perp}$ and by writing Q as in (4.71), eqn(4.74a) is reduced to the following equivalent set of conditions

$$FQ_\infty = GQ_\infty J_{q_\infty}(\infty),\; FQ_\beta J_{p_\beta}(\beta) = GQ_\beta,\; \forall\; \beta\in\{\alpha_1,\dots,\alpha_\mu\} \qquad (4.75c)$$

However, by conditions (4.73a) and (4.73b) it is obvious that conditions (4.75c) are automatically satisfied and this proves the sufficiency.

If we now denote by $T_w^\perp = \left[\begin{array}{c}\tilde{F}_w\\ \hline -\tilde{G}_w\end{array}\right]$ a right inverse of T_w, then by condition (4.74b), eqn(4.72) is derived. □

The matrices $\hat{R}$, defined for a given Q, which satisfies the conditions of Corollary (4.9), are not uniquely defined, since the right inverse is not uniquely defined. The parametrisation of the family of the $\hat{R}$ matrices is described next. Before we state the result we introduce some useful notation. Let $Q=[Q_\infty,Q_{\alpha_1},\dots,Q_\beta,\dots,Q_{\alpha_\mu}]$ be n×n nonsingular matrix, where the columns of Q_β, $\forall\; \beta\in\mathcal{R}$, form a normal complete prime set of chains of (F,G) at s=β. If Q_β is of dimension $n\times p_\beta$ and $\mathcal{I}_\beta=\{(d_1,\sigma_1),\dots,(d_\rho,\sigma_\rho);\; d_1<\dots<d_\rho\}$ is the index set of (F,G) at s=β, $p_\beta=\sum_{i=1}^{\rho} d_i$, then we define as the $J_{p_\beta}(\beta)$ matrix of Q_β, for $\beta\in\{\alpha_1,\dots,\alpha_\mu\}$, the matrix

$$J_{p_\beta} = \text{diag}\{\underbrace{J_{d_1}(\beta);\dots;J_{d_1}(\beta)}_{\sigma_1};\dots;\underbrace{J_{d_\rho}(\beta);\dots;J_{d_\rho}(\beta)}_{\sigma_\rho}\} \qquad (4.76a)$$

where $J_{d_i}(\beta)$ are standard Jordan blocks for β of dimension d_i. For the case of β=∞, and if $\mathcal{I}_\infty=\{(q_1,t_1),\dots,(q_\nu,t_\nu);q_1<\dots<q_\nu\}$ is the index set of (F,G) at s=∞, $q_\infty=\sum_{i=1}^{\nu} q_i$, then

$$J_{q_\infty}(\infty) = \text{diag}\{\underbrace{J_{q_1}(0);\dots;J_{q_1}(0)}_{t_1};\dots;\underbrace{J_{q_\nu}(0);\dots;J_{q_\nu}(0)}_{t_\nu}\} \qquad (4.76b)$$

Remark (4.14): For every matrix $Q=[Q_\infty,Q_{\alpha_1},\dots,Q_\beta,\dots,Q_{\alpha_\mu}]$, where the columns of Q_β form a normal complete set of chains of (F,G) at s=β, $\beta\in\mathcal{R}$, the following properties hold true:

$$GQ_{\beta} = FQ_{\beta}J_{p_{\beta}}(\beta) \text{ for } \forall\, \beta\epsilon\{\alpha_1,\ldots,\alpha_{\mu}\} \tag{4.77a}$$

$$FQ_{\infty} = GQ_{\infty}J_{q_{\infty}}(\infty)$$

The family of matrices $\hat{R}$ is defined by.

Corollary (4.10): Let $Q=[Q_{\infty},Q_{\alpha_1},\ldots,Q_{\beta},\ldots,Q_{\alpha_{\mu}}]$ be any n×n matrix for $\forall\,\beta\epsilon R$ the columns of Q_{β} form a normal complete set of chains of (F,G) at $s=\beta$. Then,

(i) A particular solution of eqn(4.70) is given by

$$\hat{R}_o = F\underbrace{[0,Q_{\alpha_1},\ldots,Q_{\alpha_{\mu}}]}_{n\times n}+G\underbrace{[Q_{\infty},0,\ldots,0]}_{n\times n} \tag{4.78}$$

(ii) The general family of solutions of eqn(4.70) is defined by

$$\hat{R} = \hat{R}_o+F[Q_{\infty},Q_{\alpha_1}J_{p_{\alpha_1}}(\alpha_1),\ldots,Q_{\alpha_{\mu}}J_{p_{\alpha_{\mu}}}(\alpha_{\mu})]U+$$
$$-G[Q_{\infty}J_{q_{\infty}}(\infty),Q_{\alpha_1},\ldots,Q_{\alpha_{\mu}}]U \tag{4.78}$$

where U is an arbitrary n×n matrix.

Proof

(i) By inspection of (4.57a) a right inverse of T_w is defined by

$$T_w^{\dagger} = \begin{bmatrix} \tilde{F}_w \\ -\tilde{G}_w \end{bmatrix},\ \tilde{F}_w = \text{diag}\{0,I_{p_{\alpha_1}},\ldots,I_{p_{\alpha_{\mu}}}\},\ \tilde{G}_w = \text{diag}\{I_{q_{\infty}},0,\ldots,0\} \tag{4.79a}$$

and thus, by (4.72) and the partitioned form of Q (4.78) is established.

(ii) It is known ([Ra&Mit - 1]) that a general family of right inverses for T_w is given by

$$T_w^{\dagger} = \begin{bmatrix} \tilde{F}_w \\ -\tilde{G}_w \end{bmatrix} + \begin{bmatrix} F'_w \\ -G'_w \end{bmatrix}U,\ U\epsilon\mathbb{C}^{n\times n} \text{ arbitrary} \tag{4.79b}$$

where $\tilde{F}_w,-\tilde{G}_w$ define a particular right inverse (as in (4.79a)) and $\begin{bmatrix} F'_w \\ -G'_w \end{bmatrix}$ is a right annihilator of T_w. By choosing F'_w,G'_w as in (4.75b) and by noting that $\hat{R}_o=FQ\tilde{F}_w+GQ\tilde{G}_w$, (4.78) is established. □

With Corollaries (4.9) and (4.10) the problem of defining the pairs (R,Q) is solved. The general procedure for the construction of the pairs is summarised below.

Integrated nested basis matrix approach: For the n×n regular pencil sF-G, the Weierstrass form sF_w-G_w and the pairs (R,Q) for which $R(sF-G)Q=sF_w-G_w$ may be constructed as it is discussed below:

(i) As with the Segre characteristic based approach, define the root range of (F,G), $\mathcal{R}=\{\infty,\alpha_1,\dots,\alpha_\mu\}$, the indices of annihilation τ_β of (F,G) at $s=\beta$ for $\forall\ \beta\in\mathcal{R}$ and a right annihilator of $P_\beta^{\tau_\beta}(F,G)$, $\hat{P}_\beta^{\tau_\beta}$ of $P_\beta^{\tau_\beta}(F,G)$ for $\forall\ \beta\in\mathcal{R}$.

(ii) By the Segré characteristic based approach, the canonical pair (F_w,G_w) may be constructed. Alternatively, (F_w,G_w) may be found as it is described below; the following approach also yields the pairs (R,Q).

(iii) Reduce every matrix $\hat{P}_\beta^{\tau_\beta}$ to a nested basis matrix $N_\beta^{\tau_\beta}$, for every $\beta\in\mathcal{R}$, by elementary column operations.

(iv) Follow the steps of the nested basis matrix approach, described before to compute a normal complete prime set of chains $\tilde{\Sigma}_\beta(F,G)$ for $\forall\ \beta\in\mathcal{R}$. The procedure generates (F_w,G_w) in an alternative way.

(v) Construct an n×n, nonsingular matrix Q, as $Q=[Q_\infty,Q_{\alpha_1},\dots,Q_{\alpha_\mu}]$, where the columns of every block Q_β, $\beta\in\mathcal{R}$, are the vectors in the ordered set $\tilde{\Sigma}_\beta(F,G)$. The general family of $\hat{R}=R^{-1}$ matrices is then given by (4.78).

4.6 Conclusions

Three different approaches for the derivation of the Weierstrass canonical form of a regular pencil have been presented and a systematic procedure for the derivation of the pairs of transformations (R,Q) which reduce a pair (F,G) to its Weierstrass canonical description has been given. The approaches discussed present the set of strict equivalent invariants of a regular pencil as numerical invariants of the ordered pair (F,G); this provides an alternative interpretation of the e.d. of a matrix pencil to that originally given in terms of the Smith form. The advantage of the approach is that it highlights the geometric aspects of the set of e.d., as these are defined by the properties of the nested basis matrices and the structure of the maximal generalised nullspaces. The geometry of the e.d. will be considered in some more details in Chapter 7; the results presented here will form the basis for the study of more general properties of the geometry of matrix pencils. The central feature of the present approach is that it is based on the rank properties and nullspaces properties of α-Toeplitz matrices; thus, an efficient singular value decomposition algorithm for such matrices is the only numerical tool needed for the computations of (F_w, G_w) and of the pairs (R,Q). The aim of the following chapter is to extend the present approach to the case of singular pencils.

CHAPTER 5:

Number Theoretic and Geometric aspects of the column and row minimal indices of a Singular Pencil

CHAPTER 5: NUMBER THEORETIC AND GEOMETRIC ASPECTS OF THE COLUMN AND ROW MINIMAL INDICES OF A SINGULAR PENCIL

5.1 Introduction

The aim of this chapter is to extend the number theoretic results, derived in chapter (4) for the characterisation of the e.d. structure of a regular pencil, to the case of c.m.i. and r.m.i. of a singular pencil and to provide a detailed study of the geometry of subspaces associated with the c.m.i., r.m.i. of a singular pencil. The number theoretic properties of c.m.i. (r.m.i.) are shown to be similar to those of an e.d. at $s=\alpha$, but the results are now based on the Toeplitz matrices of the pair (F,G) and not on the truncated α- (∞-) Toeplitz matrices used for the e.d. of a regular pencil. The unifying property between the present treatment and that of chapter (4) is the use of Piecewise Arithmetic Progression sequences. The geometric aspects of c.m.i. (r.m.i.) stem from the properties of the subspaces associated with the homogeneous polynomial vectors characterising the set of c.m.i. (r.m.i.). Such vectors are defined from the vectors in the right (left) nullspaces of the appropriate Toeplitz matrices and thus their definition is geometric, rather than algebraic (the classical approach is based on the theory of minimal bases of rational vector spaces).

The results presented here provide: First, an alternative procedure for the computation of minimal indices, which is independent from the use of strict equivalence transformations ([Gant. -1], [Van Do. -1]) and independent from the algebraic minimal basis approach ([For. -1]). Second, a purely geometric approach to the minimal basis theory [For. 1] . The geometric aspects of a singular pencil emerge as byproducts of the properties of Teoplitz matrices and thus they may be discussed as properties of the ordered pair (F,G) and independently from the pencil $sF-\hat{s}G$.

5.2 Toeplitz matrices and characteristic spaces of (F,G): definitions and preliminary results

In this section the notion of right and left Toeplitz matrices which may be associated with a general pencil $sF-\hat{s}G$, or a pair (F,G) is introduced and some preliminary nature results on the right, left nullspaces correspondingly are derived. These results form the basis for the analysis presented in the following section. The pencil $sF-\hat{s}G$ (or the pair (F,G)), is assumed to be a general singular pencil (pair).

Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ and let $\text{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}=\rho\leq\min(m,n)$. For singular pencils, a complete set of invariants under strict equivalence is defined by the set of e. d. (finite and infinite) and the sets of c.m.i. and r.m.i. [Gan. -1], [Tur. & Ait. -1]. The minimal indices arise because of the singularity of the pencil, which in turn implies linear dependence amongst its columns and/or its rows; thus there exist polynomial vectors $\underline{x}(s,\hat{s}) \in \mathbb{R}^{n}[s,\hat{s}]$, $\underline{y}(s,\hat{s}) \in \mathbb{R}^{m}[s,\hat{s}]$ such that at least one of the following conditions is satisfied

$$\{sF-\hat{s}G\}\underline{x}(s,\hat{s})=\underline{0} \iff \underline{x}(s,\hat{s})\in N_{r}\{sF-\hat{s}G\} \tag{5.1}$$

$$\underline{y}^{t}(s,\hat{s})\{sF-\hat{s}G\}=\underline{0} \iff \underline{y}^{t}(s,\hat{s})\in N_{\ell}\{sF-\hat{s}G\} \tag{5.2}$$

where $N_{r}\{sF-\hat{s}G\}$, $N_{\ell}\{sF-\hat{s}G\}$ denote the right, left nullspaces over $\mathbb{R}(s,\hat{s})$ (binary rational vector spaces) respectively of $sF-\hat{s}G$. The binary vectors $\underline{x}(s,\hat{s})$ and $\underline{y}^{t}(s,\hat{s})$ express dependence relationships among the columns and rows correspondingly of $sF-\hat{s}G$; conversely, every such relationship may be represented by a binary polynomial vector. Given that a non-homogeneous identity in $(s,\hat{s})$ implies that each of its distinct homogeneous parts is an identity, there is no loss of generality in confining the discussion to the homogeneous identity of any order; thus, homogeneous polynomials $\underline{x}(s,\hat{s})$ and $\underline{y}^{t}(s,\hat{s})$ are considered. Nonhomogeneous solutions may always be expressed in terms of fundamental homogeneous solutions.

Let $\mathbb{R}_d[s,\hat{s}]$ be the abelian additive group of homogeneous polynomials with degree d and let $\mathbb{R}_d[s,\hat{s}]^n$ be the n-vector (n-tuple) with elements from $\mathbb{R}_d[s,\hat{s}]$. If $\underline{x}(s,\hat{s})\in\mathbb{R}_d[s,\hat{s}]^n$, $\underline{y}(s,\hat{s})\in\mathbb{R}_c^m[s,\hat{s}]$, then we may write

$$\underline{x}(s,\hat{s})=\underline{x}_{o,d}\hat{s}^d+\underline{x}_{1,d-1}s\hat{s}^{d-1}+\dots+\underline{x}_{d-1,1}s^{d-1}\hat{s}+\underline{x}_{d,o}s^d=X_d\underline{e}_d(s,\hat{s}) \tag{5.3a}$$

$$\underline{y}^t(s,\hat{s})=\underline{y}^t_{o,c}\hat{s}^c+\underline{y}^t_{1,c-1}s\hat{s}^{c-1}+\dots+\underline{y}^t_{c-1,1}s^{c-1}\hat{s}+\underline{y}^t_{c,o}s^c=\underline{e}^t_c(s,\hat{s})Y_c \tag{5.3b}$$

where $\underline{e}_k(s,\hat{s})=[\hat{s}^k,s\hat{s}^{k-1},\dots,s^{k-1}\hat{s},s^k]^t$. The matrices $X_d\in\mathbb{R}^{n\times(d+1)}$, $Y_c\in\mathbb{R}^{(c+1)\times m}$, uniquely characterise the homogeneous binary polynomial vectors $\underline{x}(s,\hat{s})$, $\underline{y}^t(s,\hat{s})$ correspondingly and shall be referred to as <u>basis matrices</u> of $\underline{x}(s,\hat{s})$, $\underline{y}^t(s,\hat{s})$ respectively. The real vector spaces $\mathcal{X}$=col. $\text{span}_{\mathbb{R}}\{X_d\}$ and $\mathcal{Y}$=row $\text{span}_{\mathbb{R}}\{Y_c\}$ will be called the <u>supporting spaces</u> of $\underline{x}(s,\hat{s})$, $\underline{y}^t(s,\hat{s})$ correspondingly. The vector $\underline{e}_k(s,\hat{s})$ will be referred to as the <u>k-th vector of apolarity</u> [Tur. & Ait. -1]. The characterisation of the binary homogeneous polynomial vectors which satisfy conditions (5.1) or (5.2) is given by the following result.

<u>Proposition (5.1)</u>: Let $\underline{x}(s,\hat{s})=X_d\underline{e}_d(s,\hat{s})\in\mathbb{R}_d[s,\hat{s}]^n$, $\underline{y}^t(s,\hat{s})=\underline{e}^t_c(s,\hat{s})Y_d\in\mathbb{R}_c[s,\hat{s}]^{1\times m}$, where X_d,Y_d are defined as in (5.3a),(5.3b).

(i) The condition $\{sF-\hat{s}G\}\underline{x}(s,\hat{s})=0$ is equivalent to

$$F\underline{x}_{d,o}=\underline{0},F\underline{x}_{d-1,1}=G\underline{x}_{d,o},\dots,F\underline{x}_{o,d}=G\underline{x}_{1,d-1},\underline{0}=G\underline{x}_{o,d} \tag{5.4a}$$

or equivalently,

$$\begin{bmatrix} F & 0 & \dots & 0 & 0 \\ -G & F & \dots & 0 & 0 \\ & & & & \\ 0 & 0 & \dots & -G & F \\ 0 & 0 & \dots & 0 & -G \end{bmatrix}\begin{bmatrix} \underline{x}_{d,o} \\ \underline{x}_{d-1,1} \\ \\ \underline{x}_{1,d-1} \\ \underline{x}_{o,d} \end{bmatrix}=\underline{0} \tag{5.4b}$$

(ii) The condition $\underline{y}^t(s,\hat{s})\{sF-\hat{s}G\}=\underline{0}^t$ is equivalent to

$$\underline{y}^t_{c,o}F=\underline{0}^t,\underline{y}^t_{c-1,1}F=\underline{y}^t_{c,o}G,\dots,\underline{y}^t_{o,c}F=\underline{y}^t_{1,c-1}G,\underline{0}^t=\underline{y}^t_{o,c}G \tag{5.5a}$$

or equivalently

$$[\underline{y}^t_{c,o},\underline{y}^t_{c-1,1},\dots,\underline{y}^t_{1,c-1},\underline{y}^t_{o,c}]\begin{bmatrix} F & -G & \dots & 0 & 0 \\ 0 & F & \dots & 0 & 0 \\ & & & & \\ 0 & 0 & \dots & -G & 0 \\ 0 & 0 & \dots & F & -G \end{bmatrix} = \underline{0}^t \qquad (5.5b)$$ □

The proof of this result follows by substituting $\underline{x}(s,\hat{s}),\underline{y}^t(s,\hat{s})$, expressed as in (5.3a),(5.3b), into (5.1),(5.2) correspondingly and by equating coefficients of equal powers. Condition (5.4b),(5.5b) imply that the study of homogeneous binary vectors in $N_r\{sF-\hat{s}G\},N_\ell\{sF-\hat{s}G\}$ may be reduced to a study of right, left nullspaces of matrices defined from the ordered pair (F,G). The matrices defined by

$$T_1(F,G)\triangleq\begin{bmatrix} F \\ -G \end{bmatrix}, T_2(F,G)\triangleq\begin{bmatrix} F & 0 \\ -G & F \\ 0 & -G \end{bmatrix},\dots,T_k(F,G)\triangleq\underset{\longleftarrow\; k \;\text{blocks}\;\longrightarrow}{\begin{bmatrix} F & 0 & \dots & 0 & 0 \\ -G & F & \dots & 0 & 0 \\ & & & & \\ 0 & 0 & \dots & -G & F \\ 0 & 0 & \dots & 0 & -G \end{bmatrix}}\begin{matrix}\uparrow \\ k+1 \\ \text{blocks} \\ \downarrow\end{matrix} \qquad (5.6)$$

$$\tilde{T}_1(F,G)\triangleq[F,-G],\tilde{T}_2(F,G)\triangleq\begin{bmatrix} F & -G & 0 \\ 0 & F & -G \end{bmatrix},\dots,\tilde{T}_k(F,G)\triangleq\underset{\longleftarrow\; k+1 \;\text{blocks}\;\longrightarrow}{\begin{bmatrix} F & -G & \dots & 0 & 0 \\ 0 & F & \dots & 0 & 0 \\ & & & & \\ 0 & 0 & \dots & -G & 0 \\ 0 & 0 & \dots & F & -G \end{bmatrix}}\begin{matrix}\uparrow \\ k \\ \text{blocks} \\ \downarrow\end{matrix} \qquad (5.7)$$

are clearly Toeplitz matrices; $T_k(F,G)\in\mathbb{R}^{(k+1)m\times kn}$, $\tilde{T}_k(F,G)\in\mathbb{R}^{km\times(k+1)n}$ and shall be called <u>k-th order right-</u>,<u>left Toeplitz matrices</u> of (F,G) respectively. The vector spaces $N^k_r\triangleq N_r\{T_k(F,G)\},N^k_\ell\triangleq N_\ell\{\tilde{T}_k(F,G)\}$ will be referred to as the <u>k-th right-</u>,<u>left characteristic spaces</u> of (F,G) correspondingly; vectors $\underline{x}_k\in\mathbb{R}^{kn},\underline{x}_k\in N^k_r$ and $\underline{y}_k\in\mathbb{R}^{km},\underline{y}_k\in N^k_\ell$ will be called <u>k-th right-</u> and <u>k-th left-annihilating vectors</u> of (F,G) respectively. Every k-th right annihilating vector $\underline{x}_k$ and k-th left annihilating vector $\underline{y}^t_k$ may be partitioned as

$$\underline{x}_k = [\underline{x}^t_{k-1,o},\underline{x}^t_{k-2,1},\dots,\underline{x}^t_{1,k-2},\underline{x}^t_{o,k-1}]^t \tag{5.5a}$$

$$\underline{y}^t_k = [\underline{y}^t_{k-1,o},\underline{y}^t_{k-2,1},\dots,\underline{y}^t_{1,k-2},\underline{y}^t_{0,k-1}] \tag{5.5b}$$

and the binary polynomial vectors defined from $\underline{x}_k,\underline{y}^t_k$ by

$$\underline{x}(\underline{x}_k;s,\hat{s}) = [\underline{x}_{o,k-1},\underline{x}_{1,k-2},\dots,\underline{x}_{k-2,1},\underline{x}_{k-1,o}]\underline{e}_{k-1}(s,\hat{s}) = X_{k-1}\underline{e}_{k-1}(s,\hat{s}) \tag{5.6a}$$

$$\underline{y}^t(\underline{y}^t_k;s,\hat{s}) = \underline{e}^t_{k-1}(s,\hat{s})\begin{bmatrix}\underline{y}^t_{o,k-1}\\ \underline{y}^t_{1,k-2}\\ \vdots\\ \underline{y}^t_{k-2,1}\\ \underline{y}^t_{k-1,o}\end{bmatrix} = \underline{e}^t_{k-1}(s,\hat{s})Y_{k-1} \tag{5.6b}$$

will be referred to as the associated <u>k-th right-</u>, <u>left-annihilating polynomial vectors</u> generated by $\underline{x}_k,\underline{y}^t_k$ correspondingly. The supporting subspaces $\mathcal{X}(\underline{x}_k),\mathcal{Y}(\underline{y}^t_k)$ of the annihilating polynomial vectors $\underline{x}(\underline{x}_k;s,\hat{s})$, $\underline{y}^t(\underline{y}^t_k;s,\hat{s})$ will be referred to as the <u>associated k-th right-</u>, <u>left-annihilating spaces</u> of $\underline{x}_k,\underline{y}^t_k$ respectively.

There is an obvious duality between the results concerning the structure of right and left Toeplitz matrices of (F,G); the basis of this duality is the fact that $\tilde{T}_k(F,G)^t = T_k(F,G)$. Thus, results concerning $T_k(F,G)$ may be translated into corresponding results for $\tilde{T}_k(F,G)$ and vice-versa. In the following, the case of right Toeplitz matrices will be considered and the interpretation of the results to left Toeplitz matrices is rather obvious. A useful property of annihilating vectors, which readily follows from Proposition (5.1) is stated below.

<u>Remark (5.1)</u>: Let $\underline{x}_k$ be a k-th right annihilating vector of (F,G). Every vector $\underline{x}_r$ derived by a trivial expansion of $\underline{x}_k$ as $\underline{x}_r = [\underline{0}^t;\dots;\underline{0}^t;\underline{x}^t_k;\underline{0}^t;\dots\underline{0}^t]$ $\in \mathbb{R}^{rn}$, where $\underline{0}$ is an n-dimensional zero vector, is an r-th right annihilating vector of (F,G).

The Toeplitz matrices of (F,G) have been used by Gantmacher [Gan. -1] for the definition of c.m.i. and r.m.i. of a pencil. The procedure suggested in [Gan. -1] for determining the c.m.i. involves the following steps: Find the smallest integer ε for which $N_r^{\varepsilon} \neq \{0\}$; then ε is the value of the smallest c.m.i. By strict equivalence transformation reduce $sF-\hat{s}G$ to the pencil

$$sF'-\hat{s}G' = \begin{bmatrix} L_{\varepsilon}(s,\hat{s}) & 0 \\ 0 & s\tilde{F}-\hat{s}\tilde{G} \end{bmatrix} \tag{5.8}$$

where $L_{\varepsilon}(s,\hat{s})$ is a standard c.m.i. block associated with ε. The pencil $s\tilde{F}-\hat{s}\tilde{G}$ is considered next and the procedure is repeated. The main objective here is to study the structure of the Toeplitz matrices and find the set of c.m.i. without having to resort to the use of strict equivalence transformations implied in the above procedure. The dominant idea in the present study is that the set of N_r^k spaces contains all the information needed to find the set of c.m.i. Our first step in our study of the right characteristic spaces is the determination of their dimension. The following standard result will be used.

Lemma (5.1) [Gan. -1]: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$, $\rho = \text{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}$. There exists a pair (R,Q) of strict equivalence transformations such that $sF-\hat{s}G \xrightarrow{(R,Q)} R(sF-\hat{s}G)Q = sF'-\hat{s}G'$, where

$$sF'-\hat{s}G' = \left[\begin{array}{c:c:c} sF_c-\hat{s}G_c & & \\ \hdashline & sF_r-\hat{s}G_r & \\ \hdashline & & sA-\hat{s}B \end{array}\right] \tag{5.9}$$

$sA-\hat{s}B$ is regular, $sF_c-\hat{s}G_c$ is characterised by c.m.i. and $sF_r-\hat{s}G_r$ is characterised by r.m.i. only. □

The pencils $sA-\hat{s}B, sF_c-\hat{s}G_c, sF_r-\hat{s}G_r$, defined for a general singular pencil $sF-\hat{s}G$, will be referred to as a <u>regular-</u>, <u>right-</u>, <u>left-restrictions</u> correspondingly of $sF-\hat{s}G$. A singular pencil characterised only by r.m.i.

will be called entirely left singular; if it is characterised only by c.m.i., then it will be called entirely right singular. The restrictions $sA-\hat{s}B, sF_c-\hat{s}G_c, sF_r-\hat{s}G_r$ of a pencil $sF-\hat{s}G$ are not uniquely defined; the following property, however, is readily verified.

Remark (5.2): If $\{sA-\hat{s}B, sF_c-\hat{s}G_c, sF_r-\hat{s}G_r\}, \{sA'-\hat{s}B', sF_c'-\hat{s}G_c', sF_r'-\hat{s}G_r'\}$ are two triples of a regular-, right-, left-strictrictions of $sF-\hat{s}G$, then $(sA-\hat{s}B)\mathcal{E}_s(sA'-\hat{s}B'), (sF_c-\hat{s}G_c)\mathcal{E}_s(sF_c'-\hat{s}G_c'), (sF_r-\hat{s}G_r)\mathcal{E}_s(sF_r'-\hat{s}G_r')$.

For regular, entirely left singular, and entirely right singular pencils we have the following properties:

Proposition (5.2): Let $sF-\hat{s}G$ be a pencil and let $\mathcal{N}_r^k(F,G), \mathcal{N}_\ell^k(F,G)$ be the corresponding k-th right, left characteristic spaces of (F,G). The following properties hold true:

(i) $sF-\hat{s}G$ is regular if and only if $\mathcal{N}_r^k(F,G)=\{0\}, \mathcal{N}_\ell^k(F,G)=\{0\}$ for all k, $k=1,2,\ldots$

(ii) If $sF-\hat{s}G$ is entirely left singular, then $\mathcal{N}_r^k(F,G)=\{0\}$ for all k, $k=1,2,\ldots$

(iii) If $sF-\hat{s}G$ is entirely right singular, then $\mathcal{N}_\ell^k(F,G)=\{0\}$ for all k, $k=1,2.$

Proof

(i) If $sF-\hat{s}G$ is regular, then $\mathcal{N}_r\{sF-\hat{s}G\}=\{0\}, \mathcal{N}_\ell\{sF-\hat{s}G\}=\{0\}$ and thus there is no nonzero vectors $\underline{x}(s,\hat{s}) \in \mathcal{N}_r\{sF-\hat{s}G\}, \underline{y}^t(s,\hat{s}) \in \mathcal{N}_\ell\{sF-\hat{s}G\}$; these two conditions imply that the only solutions of equations

$$T_k(F,G)\underline{x}_k = 0, \quad \underline{y}_k^t\tilde{T}_k(F,G) = 0 \quad \text{for all } k=1,2,\ldots \tag{5.10}$$

are $\underline{x}_k=0, \underline{y}_k=0$. Conversely, if (5.10) hold true for all k, then the only vectors $\underline{x}(s,\hat{s}), \underline{y}^t(s,\hat{s})$ which may be found such that (5.1) and (5.2) are satisfied, are the zero vectors. This clearly implies that $sF-\hat{s}G$ is regular.

(ii) Assume that $N_r^k(F,G)\neq\{0\}$. Then, there exists a k-th right annihilating vector $\underline{x}_k$ and thus an annihilating polynomial vector $\underline{x}(\underline{x}_k;s,\hat{s})$ such that $(sF-\hat{s}G)\underline{x}(\underline{x}_k;s,\hat{s})=0$. This condition implies that $N_r\{sF-\hat{s}G\}\neq\{0\}$ and that $sF-\hat{s}G$ is characterised also by c.m.i.; this contradicts the assumption that the pencil is entirely left singular. The proof of part (iii) is identical.

□

The importance of entirely right (left) singular pencils in the study of right (left) characteristic spaces of a general pencil is demonstrated by the following result.

<u>Proposition (5.3)</u>: Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ and let $sF_c-\hat{s}G_c$ be a right restriction of $sF-\hat{s}G$. If $N_r^k(F,G), N_r^k(F_c,G_c)$ are the k-th right characteristic spaces of $(F,G),(F_c,G_c)$ correspondingly, then

$$\dim N_r^k(F,G) = \dim N_r^k(F_c,G_c) \quad \text{for } \forall\, k,\ k=1,2,\ldots$$

<u>Proof</u>

Let $\underline{x}_k\in N_r^k(F,G), \underline{x}_k\neq 0$, and let us partition $\underline{x}_k$ according to the partitioning of $T_k(F,G)$, as $\underline{x}_k=[\underline{x}_{1,k}^t,\ldots,\underline{x}_{i,k}^t,\ldots,\underline{x}_{k,k}^t]^t$, where $\underline{x}_{i,k}\in\mathbb{R}^n$, $\forall\, i\in\underset{\sim}{k}$. Then $T_k(F,G)\underline{x}_k=0$ implies that

$$F\underline{x}_{1,k}=0, G\underline{x}_{1,k}=F\underline{x}_{2,k},\ldots,G\underline{x}_{k-1,k}=F\underline{x}_{k,k}, G\underline{x}_{k,k}=0 \qquad (5.11)$$

Let (R,Q) be a pair of strict equivalence transformations which reduce $sF-\hat{s}G$ to the decomposition (5.9). By writing $\underline{x}_{i,k}=Q\underline{x}'_{i,k}$, $\forall\, i\in\underset{\sim}{k}$, and premultiplying eqn(5.11) by R we obtain the equivalent conditions

$$F'\underline{x}'_{1,k}=0, G'\underline{x}'_{1,k}=F\underline{x}'_{2,k},\ldots,G'\underline{x}'_{k-1,k}=F'\underline{x}'_{k,k}, G'\underline{x}'_{k,k}=0 \qquad (5.12a)$$

since $sF'-\hat{s}G'=\text{block diag}\{sF_c-\hat{s}G_c; sF_r-\hat{s}G_r; sA-\hat{s}B\}$, we may partition every $\underline{x}'_{i,k}$, $\forall k$, according to the partitioning of $sF'-\hat{s}G'$ as

$$\underline{x}'_{i,k} = [\underline{x}_{i,k}^{c\,t}, \underline{x}_{i,k}^{r\,t}, \underline{x}_{i,k}^{w\,t}]^t \qquad (5.12b)$$

By using the partitioning of $\underline{x}^{\prime}_{i,k}$ as above, equations (5.12a) are reduced to the following equivalent set

$$F_c\underline{x}^c_{1,k}=\underline{0},G_c\underline{x}^c_{1,k}=F_c\underline{x}^c_{2,k},\ldots,G_c\underline{x}^c_{k-1,k}=F_c\underline{x}^c_{k,k},G_c\underline{x}^c_{k,k}=\underline{0} \qquad (5.13a)$$

$$F_r\underline{x}^r_{1,k}=0,G_r\underline{x}^r_{1,k}=F_r\underline{x}^r_{2,k},\ldots,G_r\underline{x}^r_{k-1,k}=F_r\underline{x}^r_{k,k},G_r\underline{x}^r_{k,k}=\underline{0} \qquad (5.13b)$$

$$A\underline{x}^w_{1,k}=0,B\underline{x}^w_{1,k}=A\underline{x}^w_{2,k},\ldots,B\underline{x}^w_{k-1,k}=A\underline{x}^w_{k,k},B\underline{x}^w_{k,k}=\underline{0} \qquad (5.13c)$$

The pencil $sF_r-\hat{s}G_r$ is entirely left singular and $sA-\hat{s}B$ is regular; thus, by Proposition (5.2) conditions (5.13b) and (5.13c) have as the only possible solution the zero vectors, i.e. $\underline{x}^r_{i,k}=\underline{0},\underline{x}^w_{i,k}=\underline{0}$ for $\forall\, k, k=1,2,\ldots$ and $\forall\, i, i\in\underset{\sim}{k}$. We may thus write that $\underline{x}^{\prime}_{i,k}=[\underline{x}^{c\;t}_{i,k};\underline{0}^t;\underline{0}^t]^t$ and it is obvious that the number of independent vectors in $\mathcal{N}^k_r(F,G)$ is equal to the number of independent vectors in $\mathcal{N}^k_r(F_c,G_c)$. □

This result demonstrates that the study of $\mathcal{N}^k_r(F,G)$ is reduced to a study of $\mathcal{N}^k_r(F_c,G_c)$, where $sF_c-\hat{s}G_c$ is a right restriction of $sF-\hat{s}G$ and thus an entirely right singular pencil. An obvious Corollary of the above result is stated below.

<u>Corollary (5.1)</u>: Let $sF-\hat{s}G$ be an $m\times n$ singular pencil and let $sF_c-\hat{s}G_c$ be an $r\times p$ right restriction of $sF-\hat{s}G$. If N^k_c is a basis matrix of $\mathcal{N}^k_r(F_c,G_c)$, then there exists a $Q\in\mathbb{R}^{n\times n}, |Q|\neq 0$ such that a basis matrix N^k for $\mathcal{N}^k_r(F,G)$ may be expressed in terms of N^k_c as

$$N^k=\mathrm{diag}\{\underbrace{Q,\ldots,Q}_{k}\}\begin{bmatrix}\begin{bmatrix}X_{1,k}\\ \hdashline 0\end{bmatrix}\\ \hdashline \begin{bmatrix}X_{2,k}\\ \hdashline 0\end{bmatrix}\\ \hdashline \vdots \\ \hdashline \begin{bmatrix}X_{k,k}\\ \hdashline 0\end{bmatrix}\end{bmatrix}\begin{matrix}\updownarrow p\\ \updownarrow n-p\\ \updownarrow p\\ \updownarrow n-p\\ \\ \updownarrow p\\ \updownarrow n-p\end{matrix}\;,\text{ where } N^k_c=\begin{bmatrix}X_{1,k}\\ \hdashline X_{2,k}\\ \hdashline \vdots\\ \hdashline X_{k,k}\end{bmatrix}\begin{matrix}\updownarrow p\\ \updownarrow p\\ \vdots\\ \updownarrow p\end{matrix} \qquad (5.14)$$

□

In the following, attention is focussed on the study of properties of the right characteristic space of entirely right singular matrix pencils. The derived results form the basis for the study of the right characteristic spaces of a general pencil. Note that by duality, we may obtain similar results for the left characteristic spaces of entirely left singular pencils, which are essential for the study of left characteristic spaces of a general pencil.

5.3 The right characteristic spaces of entirely right singular pencils

Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be an entirely right singular pencil and let $I_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}$ be the set of c.m.i. of $sF-\hat{s}G$. The set $I_c(F,G)$ is a complete set of invariants for the strict equivalence class of $sF-\hat{s}G$; this equivalence class has a canonical element, the Kronecker form, which is of the following type

$$sF_k-\hat{s}G_k = \left[\, \underset{\leftarrow g \rightarrow}{\bigcirc} \;\vdots\; \text{block-diag}\{L_{\varepsilon_{g+1}}(s,\hat{s});\ldots;L_{\varepsilon_p}(s,\hat{s})\} \,\right] \tag{5.15a}$$

where for $\forall \varepsilon>0$, the associated block $L_\varepsilon(s,\hat{s})$ has the following form

$$L_\varepsilon(s,\hat{s}) = sL_\varepsilon-\hat{s}\hat{L}_\varepsilon, \text{ where } L_\varepsilon = \underbrace{\left[I_\varepsilon \;\vdots\; \underline{0} \right]}_{\leftarrow \varepsilon+1 \rightarrow} \updownarrow \varepsilon, \; \hat{L}_\varepsilon = \underbrace{\left[\underline{0} \;\vdots\; I_\varepsilon \right]}_{\leftarrow \varepsilon+1 \rightarrow} \updownarrow \varepsilon \tag{5.15b}$$

The canonical pencil $L_\varepsilon(s,\hat{s})=sL_\varepsilon-\hat{s}\hat{L}_\varepsilon$ plays a key role in the study of the right characteristic space of a general pencil $sF-\hat{s}G$; we start off by giving some preliminary results on the properties of $N_r^k(L_\varepsilon,\hat{L}_\varepsilon)$.

<u>Lemma (5.2)</u>: Let $L_\varepsilon,\hat{L}_\varepsilon \in \mathbb{R}^{\varepsilon\times(\varepsilon+1)}$ be the matrices defined by (5.15b). The set of matrix equations

$$L_\varepsilon \underline{x}_1=0, \hat{L}_\varepsilon \underline{x}_1=L_\varepsilon \underline{x}_2,\ldots,\hat{L}_\varepsilon \underline{x}_{k-1}=L_\varepsilon \underline{x}_k \tag{5.16}$$

where $\underline{x}_i \in \mathbb{R}^{\varepsilon+1}, i\in \underset{\sim}{k}$, has always a solution. The general family of solutions is given by:

(i) If $k<\varepsilon+1$, then $\underline{x}_i=[0,\dots,0,a_1,a_2,\dots,a_i]^t$, $\forall\, i\in\underset{\sim}{k}$ and $a_i\in\mathbb{R}$ arbitrary.

(ii) If $k\geq\varepsilon+1$, then $\underline{x}_i=[0,\dots,0,a_1,\dots,a_i]^t$ for $\forall\, i\in\underset{\sim}{\varepsilon}$; for $i=\varepsilon+1,\dots,k$, $\underline{x}_i=[a_{i-\varepsilon},a_{i-\varepsilon+1},\dots,a_i]^t$ with $a_j\in\mathbb{R}$ arbitrary. □

The proof of this Lemma follows by inspection of the canonical structure of the $L_\varepsilon,\hat{L}_\varepsilon$ matrices. Using this Lemma we have:

Proposition (5.4): Let $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)$ be the right characteristic space of the canonical pencil $sL_\varepsilon-\hat{s}L_\varepsilon$. $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)\neq\{0\}$ if and only if $k\geq\varepsilon+1$; furthermore, if $k\geq\varepsilon+1$, then $\theta_k=\dim\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=k-\varepsilon$.

Proof

Let $\tilde{\underline{x}}_k=[\underline{x}_1^t,\underline{x}_2^t,\dots,\underline{x}_k^t]^t\in\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)$, where $\underline{x}_i\in\mathbb{R}^{\varepsilon+1}$, then the vectors $\underline{x}_i$ must satisfy the conditions

$$L_\varepsilon\underline{x}_1=0,\hat{L}_\varepsilon\underline{x}_1=L_\varepsilon\underline{x}_2,\dots,\hat{L}_\varepsilon\underline{x}_{k-1}=L_\varepsilon\underline{x}_k,\hat{L}_\varepsilon\underline{x}_k=0 \qquad (5.17)$$

If $k<\varepsilon+1$, then by Lemma (5.2) the general solution of the first k equations of (5.17) is given by $\underline{x}_i=[0,\dots,0,a_1,\dots,a_i]^t$, $\forall\, i\in\underset{\sim}{k}$ and thus $\hat{L}_\varepsilon\underline{x}_k=0$ implies that $a_1=a_2=\dots=a_k=0$. This shows that $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=\{0\}$ for all k such that $k<\varepsilon+1$. Furthermore, since $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=\{0\}$ it follows that if $k<\varepsilon+1$, then the matrix $T_k(L_\varepsilon,\hat{L}_\varepsilon)$ has full rank.

If $k\geq\varepsilon+1$, then by Lemma (5.2), the general solution of the first k equations of (5.17) are given by

$$\underline{x}_i=[0,\dots,0,a_1,\dots,a_i]^t \text{ if } i\in\underset{\sim}{\varepsilon},\ \underline{x}_i=[a_{i-\varepsilon},a_{i-\varepsilon+1},\dots,a_i]^t,\text{ for } i=\varepsilon+1,\dots,k$$

The last condition $\hat{L}_\varepsilon\underline{x}_k=\underline{0}$, then implies that $a_{k-\varepsilon+1}=a_{k-\varepsilon+2}=\dots=a_k=0$ and that the first $k-\varepsilon$ parameters $a_1,a_2,\dots,a_{k-\varepsilon}$ are arbitrary. The general expression for the vector $\tilde{\underline{x}}_k$ is thus obtained and by inspection it follows that there are $k-\varepsilon$ independent vectors in $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)$; thus $\dim\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=k-\varepsilon$. If $k\geq\varepsilon+1$, then $T_k(L_\varepsilon,\hat{L}_\varepsilon)$ has a right null space; the minimum dimension of this null space is $d_{\min}=k(\varepsilon+1)-(k+1)\varepsilon=k-\varepsilon=\dim\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)$; this shows that

$T_k(L_\varepsilon,\hat{L}_\varepsilon)$ has full rank. □

Remark (5.3): For all $\varepsilon>0$, the Toeplitz matrices $T_k(L_\varepsilon,\hat{L}_\varepsilon)$ have full rank for all $k=1,2,\ldots$

Remark (5.4): For all $k<\varepsilon+1$, then $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=\{0\}$.

Note that the vector solutions $\{\underline{x}_1,\ldots,\underline{x}_k\}$ of (5.16) are vector coefficients of a k-th right annihilating vector $\underline{x}(s,\hat{s})=\underline{x}_1 s^{k-1}+\underline{x}_2 s^{k-2}\hat{s}+\ldots+\underline{x}_{k-1}s\hat{s}^{k-2}+$ $+\underline{x}_k\hat{s}^{k-1}=[\underline{x}_k,\underline{x}_{k-1},\ldots,\underline{x}_2,\underline{x}_1]\underline{e}_{k-1}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ of $sL_\varepsilon-\hat{s}L_\varepsilon$. The basis matrix X_{k-1} has a canonical structure, and this is defined by the following result.

Corollary (5.2): Let $k\geq\varepsilon+1$ and let $\underline{x}(s,\hat{s})=[\underline{x}_k,\underline{x}_{k-1},\ldots,\underline{x}_1]\underline{e}_{k-1}(s,\hat{s})=$ $=D_{k,\varepsilon}\underline{e}_{k-1}(s,\hat{s})$ be a k-th right annihilating polynomial of $sL_\varepsilon-\hat{s}L_\varepsilon$. The basis matrix $D_{k,\varepsilon}\in\mathbb{R}^{(\varepsilon+1)\times k}$ has the following canonical structure

$$D_{k,\varepsilon}=\begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{k-\varepsilon} & 0 & \cdots & \cdots & & 0 \\ 0 & a_1 & a_2 & \cdots & a_{k-\varepsilon-1} & a_{k-\varepsilon} & 0 & & & \vdots \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ & & & 0 & a_1 & a_2 & \cdots & a_{k-\varepsilon-1} & a_{k-\varepsilon} & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & a_1 & a_2 & \cdots & a_{k-\varepsilon-1} & a_{k-\varepsilon} \end{bmatrix} \tag{5.18}$$

where the parameters $a_i\in\mathbb{R}$, but otherwise arbitrary. □

Matrices of $D_{k,\varepsilon}$ type have a canonical Toeplitz structure and shall be referred to as (k,ε)-Toeplitz matrices of $sL_\varepsilon-\hat{s}\hat{L}_\varepsilon$. The number of free parameters is $k-\varepsilon$; $D_{k,\varepsilon}$ defines the total family of k-th right annihilating vectors of the canonical pencil $sL_\varepsilon-\hat{s}\hat{L}_\varepsilon$.

Remark (5.5): Let $D_{k,\varepsilon}=[\underline{x}_k,\underline{x}_{k-1},\ldots,\underline{x}_1]$ be the general (arbitrary parameters) (k,ε)-Toeplitz matrix of $sL_\varepsilon-\hat{s}\hat{L}_\varepsilon$. The linear vector space generated by the k-th right annihilating vectors, $N_{k,\varepsilon}\overset{\Delta}{=}\{\tilde{\underline{x}}_k:\tilde{\underline{x}}_k=[\underline{x}_1^t,\dots,\underline{x}_k^t]^t\}$ has dimension k-ε.

We may now state the main result of this section.

Theorem (5.1): Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be an entirely right singular pencil and let $I_c(F,G)=\{\varepsilon_1=\dots=\varepsilon_g=0<\varepsilon_{g+1}\le\dots\le\varepsilon_p\}$ be the set of c.m.i. If $N_r^k(F,G)$ is the k-th right characteristic space of $sF-\hat{s}G$ then the following properties hold true:

(i) $N_r^k(F,G)=\{0\}$ if and only if $k<\min\{\varepsilon_j+1, j\in\underset{\sim}{p}\}$.

(ii) If $k\ge\min\{\varepsilon_j+1, j\in\underset{\sim}{p}\}$, then $N_r^k(F,G)\ne\{0\}$. In this case the dimension θ_k of $N_r^k(F,G)$ is given by

$$\theta_k = \sum_{j=1}^{p}\theta_k(\varepsilon_j) \tag{5.19}$$

where $\theta_k(\varepsilon_j)=0$ if $k<\varepsilon_j+1$ and $\theta_k(\varepsilon_j)=k-\varepsilon_j$ if $k\ge\varepsilon_j+1$.

Proof

(i) Let (R,Q) be a pair of strict equivalence transformations which reduce $sF-\hat{s}G$ to the Kronecker form (5.15a), i.e. $sF-\hat{s}G \xrightarrow{(R,Q)} R(sF-\hat{s}G)Q=$ $=sF_k-\hat{s}G_k$. If $\tilde{\underline{x}}_k=[\underline{x}_1^t,\underline{x}_2^t,\dots,\underline{x}_k^t]^t\in N_r^k(F,G)$, $\underline{x}_i\in\mathbb{R}^n$, then the vectors $\underline{x}_i$ must satisfy the conditions

$$F\underline{x}_1=0, G\underline{x}_1=F\underline{x}_2,\dots,G\underline{x}_{k-1}=F\underline{x}_k, G\underline{x}_k=0 \tag{5.20a}$$

By writing $\underline{x}_i=Q\underline{x}_i'$, $\forall\, i\in\underset{\sim}{k}$ and premultiplying (5.20a) by R we have

$$F_k\underline{x}_1'=0, G_k\underline{x}_1'=F_k\underline{x}_2',\dots,G_k\underline{x}_{k-1}'=F_k\underline{x}_k', G_k\underline{x}_k'=0 \tag{5.20b}$$

The vectors $\underline{x}_i'$ may be partitioned according to the partitioning of $sF_k-\hat{s}G_k$ as

$$\underline{x}_i' = [\underline{x}_{\varepsilon_o}^{it}, \underline{x}_{\varepsilon_{g+1}}^{it},\dots,\underline{x}_{\varepsilon_p}^{it}]^t \tag{5.20c}$$

where $\underline{x}_{\varepsilon_o}^{it}\in\mathbb{R}^g$ and corresponds to the m×g zero block of the Kronecker form and $\underline{x}_{\varepsilon_j}^{it} \in \mathbb{R}^{\varepsilon_j+1}$, $j=g+1,\dots,p$. By (5.20c) and the special structure of $sF_k-\hat{s}G_k$, conditions (5.20b) imply that the vectors $\underline{x}_{\varepsilon_o}^{i}$, $i \in \underset{\sim}{k}$ are arbitrary and that for every ε_j, $j=g+1,\dots,p$ we have the following set of conditions

$$L_{\varepsilon_j}\underline{x}_{\varepsilon_j}^{1}=0,\hat{L}_{\varepsilon_j}\underline{x}_{\varepsilon_j}^{1}=L_{\varepsilon_j}\underline{x}_{\varepsilon_j}^{2},\dots,\hat{L}_{\varepsilon_j}\underline{x}_{\varepsilon_j}^{k-1}=L_{\varepsilon}\underline{x}_{\varepsilon_j}^{k},\hat{L}_{\varepsilon_j}\underline{x}_{\varepsilon_j}^{k}=0,j=g+1,\dots,p \tag{5.20d}$$

By Proposition (5.4), we have that if $\varepsilon_j>0$, then $\mathcal{N}_r^k(L_{\varepsilon_j},\hat{L}_{\varepsilon_j})=\{0\}$, if $k<\varepsilon_j+1$ and $\dim \mathcal{N}_r^k(L_{\varepsilon_j},\hat{L}_{\varepsilon_j})=k-\varepsilon_j$, if $k \geq \varepsilon_{j+1}$. If $sF-\hat{s}G$ has no zero c.m.i. then for all $k<\min\{\varepsilon_j+1,j=g+1,\dots,p\}$, $\mathcal{N}_r^k(L_\varepsilon,\hat{L}_\varepsilon)=\{0\}$ and thus $\mathcal{N}_r^k(F,G)=\{0\}$; this condition may readily be extended to $k<\min\{\varepsilon_j+1,j \in \underset{\sim}{p}\}$ to cover the case of zero c.m.i. In fact, if a zero c.m.i. exists, then there is no k for which $\mathcal{N}_r^k(F,G)=\{0\}$; in the case there are no zero c.m.i. the condition is reduced to $k<\min\{\varepsilon_j+1,j=g+1,\dots,p\}$ and thus $\mathcal{N}_r^k(F,G)=\{0\}$. This completes the proof of part (i).

(ii) The direct sum decomposition of $\underline{x}_i'$, as in (5.20c) implies that if $\mathcal{N}_r^k(F,G)$ is nonzero, then its dimension may be found as a sum of the dimensions of the spaces generated by vectors of the type

$$\tilde{\underline{x}}_{\varepsilon_j} = [\underline{x}_{\varepsilon_j}^{1\,t},\underline{x}_{\varepsilon_j}^{2\,t},\dots,\underline{x}_{\varepsilon_j}^{k\,t}]^t \tag{5.20e}$$

where $j=0,g+1,\dots,p$. If $\varepsilon_\nu+1 \leq k<\varepsilon_{\nu+1}$ and the pencil has zero c.m.i., then we have the following: with the g zero c.m.i. there exist kg arbitrary vectors $\underline{x}_{\varepsilon_o}^{i} \in \mathbb{R}^g$ and thus the space generated by $\tilde{\underline{x}}_{\varepsilon_o}=[\underline{x}_{\varepsilon_o}^{1\,t},\dots,\underline{x}_{\varepsilon_o}^{k\,t}]^t$ has kg dimension (since there exist kg free parameters in $\tilde{\underline{x}}_{\varepsilon_o}$). By Proposition (5.4) the dimension of the space generated by $\tilde{\underline{x}}_{\varepsilon_j}$, for $\forall j$: $j=g+1,\dots,\nu$, is $k-\varepsilon_j$; by Proposition (5.4), the vectors $\tilde{\underline{x}}_{\varepsilon_j}=\underline{0}$ for j: $j=\nu+1,\dots,p$. If the pencil has no zero c.m.i., then the dimensions of $\tilde{\underline{x}}_{\varepsilon_j}$ for $j=1,\dots,\nu$ only are taken into account. □

<u>Corollary (5.3)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\ n}[s,\hat{s}]$ be an entirely right singular pencil and let $\mathcal{I}_c(F,G)$ be the set of c.m.i. Let $\varepsilon_\nu+1 \leq k<\varepsilon_{\nu+1}+1$ and $\underline{x}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ be a k-th right annihilating polynomial of $sF-\hat{s}G$. There

always exists a $Q \in \mathbb{R}^{n\times n}$, $|Q|\neq 0$ such that the family of basis matrices X_{k-1} may be expressed as

$$X_{k-1} = Q \left[\begin{array}{c} D_{k,o,g} \\ \hdashline D_{k,\varepsilon_{g+1}} \\ \hdashline \\ \\ \hdashline D_{k,\varepsilon_{\nu}} \\ \hdashline O \end{array}\right] \in \mathbb{R}^{n\times k} \qquad (5.21)$$

where D_{k,ε_j} are the (k,ε_j)-Toeplitz matrices corresponding to the nonzero c.m.i. ε_j for $j=g+1,\ldots,\nu$ and $D_{k,o,g} \in \mathbb{R}^{g\times k}$ is an arbitrary matrix associated with the g zero c.m.i. □

The proof of this result readily follows from the proof of Theorem (5.1). It is clear that the dimensions of the $N_r^k(F,G)$ spaces of an entirely right singular pencil are functions of the set of c.m.i. only. In the following, attention is focussed on the problem of determining the set of c.m.i. from the dimensions θ_k of the $N_r^k(F,G)$ spaces. From now on, we shall adopt the following representation for the set of c.m.i. of sF-ŝG: $I_c(F,G) \overset{\Delta}{=} \{(\varepsilon_i,\rho_i),\ i\in\underset{\sim}{\mu}:\ 0\leq\varepsilon_1<\ldots<\varepsilon_\mu$, where ρ_i is the multiplicity of $\varepsilon_i\}$; clearly $\{\varepsilon_i, i\in\underset{\sim}{\mu}\}$ denotes the set of distinct values of the c.m.i. of the general pencil sF-ŝG and shall be referred to as the <u>right singular set</u> of sF-ŝG. The results derived for the case of entirely right singular pencils are used next to the case of a general matrix pencil.

5.4 (F,G)-Piecewise arithmetic progression sequences and the minimal indices of a general singular pencil

The study of the dimensions of the $N_r^k(F,G)$ characteristic spaces of an entirely right singular pencil reveals that there exist strong similarities between the present study and that developed in Chapter (4) for the computation of the Segré characteristics of a regular pencil. These links will be further explored in this section and it will be shown that certain piecewise arithmetic progression sequences may be used for the computation of minimal indices. The results developed in the last section for the dimensions of $\mathcal{N}_r^k(F,G)$ spaces of an entirely right singular pencil may be readily extended to the case of a general singular pencil; this is due to Proposition (5.3), which establishes that dim $\mathcal{N}_r^k(F,G) = \dim \mathcal{N}_r^k(F_c,G_c)$, where $sF_c-\hat{s}G_c$ is a right restriction of $sF-\hat{s}G$. Matrix pencils with a right restriction will be called <u>right singular</u>, and those with a left restriction, <u>left singular</u>; clearly a pencil may be both right and left singular.

<u>Remark (5.6)</u>: A singular pencil $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ is right singular, if and only if there exists a k, k=1,2,... for which $\mathcal{N}_r^k(F,G)\neq\{0\}$. Similarly, the pencil is left singular, if and only if there exists a k, k=1,2,... for which $\mathcal{N}_\ell^k(F,G)\neq\{0\}$.

In the following, the case of right singular pencils will be considered; the results then may be interpreted to the case of left singular pencils by "transposed duality". A result characterising right (left) singularity of a pencil is considered first.

<u>Proposition (5.5)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a singular pencil. Then,

(i) $sF-\hat{s}G$ is right singular if and only if for some $k\in\{1,2,\dots,\tilde{\sigma}$: where $\tilde{\sigma}=m+1$, if $m<n$ and $\tilde{\sigma}=n$ if $m\geq n\}$ $\mathcal{N}_r^k(F,G)\neq\{0\}$.

(ii) $sF-\hat{s}G$ is left singular if and only if for some $k\in\{1,2,\dots,\tilde{\rho}$: where $\tilde{\rho}=m$, if $m\leq n$ and $\tilde{\rho}=n+1$ if $m>n\}$, $\mathcal{N}_r^k(F,G)\neq\{0\}$. □

The proof of this result is readily established by Remark (5.6) and by the maximal possible dimension of a c.m.i. (r.m.i.) block of a singular pencil. For right singular pencils Theorem (5.1) and Proposition (5.3) yield the following result.

<u>Theorem (5.2)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil and let $I_c(F,G)=\{(\varepsilon_i,\rho_i),\ i\in\underline{\mu}:\ 0\leq\varepsilon_1<\dots<\varepsilon_\mu\}$ be the set of c.m.i. If $\mathcal{N}_r^k(F,G)$ is the

k-th characteristic space of $sF-\hat{s}G$, then we have:

(i) $N^1_r(F,G)\neq\{0\}$, if and only if $\varepsilon_1=0$; then $\theta_1=\dim N^1_r(F,G)=\rho_1$.

(ii) $N^k_r(F,G)=\{0\}$, and thus $\theta_k=\dim N^k_r(F,G)=0$, if and only if $k<\varepsilon_1+1=\min\{\varepsilon_i+1, i\in\underset{\sim}{\mu}\}$.

(iii) $N^k_r(F,G)\neq\{0\}$ if and only if $k\geq\min\{\varepsilon_i+1, i\in\underset{\sim}{\mu}\}$. Then $\theta_k=\dim N^k_r(F,G)$ is defined by:

(a) $$\theta_k=\sum_{i=1}^{j}\rho_i(k-\varepsilon_i) \quad \text{if} \quad \varepsilon_j+1\leq k<\varepsilon_{j+1}+1 \tag{5.22a}$$

(b) $$\theta_k=\sum_{i=1}^{\mu}\rho_i(k-\varepsilon_i) \quad \text{if} \quad k\geq\max\{\varepsilon_i+1, i\in\underset{\sim}{\mu}\}=\varepsilon_\mu+1 \tag{5.22b}$$

□

Theorem (5.2) may be used to provide a characterisation of the c.m.i. of a right singular pencil as it is shown next. In the following we shall denote by θ_k the dimension of $N^k_r(F,G)$.

<u>Corollary (5.4)</u>: Let $I_c(F,G)=\{(\varepsilon_i,\rho_i), i\in\underset{\sim}{\mu}, 0\leq\varepsilon_1<\ldots<\varepsilon_\mu\}$ be the c.m.i. set of $sF-\hat{s}G$. The following properties hold true:

(i) $\theta_k-\theta_{k-1}=\sum_{j=1}^{t}\rho_j$, if $\varepsilon_t+1\leq k<\varepsilon_{t+1}+1$

(ii) $\theta_k-\theta_{k-1}=\sum_{j=1}^{\mu}\rho_j$, if $\varepsilon_\mu+1\leq k$

(iii) $\theta_k-\theta_{k-1}=0$, if $k<\varepsilon_1+1$, where $\theta_o=0$.

<u>Proof</u>

The result is proved by distinguishing the following cases:

(α) $\varepsilon_t+1\leq k-1<k<\varepsilon_{t+1}+1$, ($\beta$) $\varepsilon_t+1\leq k-1<k=\varepsilon_{t+1}+1$

(γ) $\varepsilon_\mu+1\leq k-1<k$, ($\delta$) $k-1<\varepsilon_1+1=k$, (ε) $k-1<k<\varepsilon_1+1$

Thus, we have:

(α) $\theta_k-\theta_{k-1}=\sum_{j=1}^{t}\rho_j(k-\varepsilon_j)-\sum_{j=1}^{t}\rho_j(k-1-\varepsilon_j)=\sum_{j=1}^{t}\rho_j$

(β) $\theta_k-\theta_{k-1}=\sum_{j=1}^{t+1}\rho_j(k-\varepsilon_j)-\sum_{j=1}^{t}\rho_j(k-1-\varepsilon_j)=\rho_{t+1}(\varepsilon_{t+1}+1-\varepsilon_{t+1})+\sum_{j=1}^{t}\rho_j=\sum_{j=1}^{t+1}\rho_j$

(γ) $\theta_k-\theta_{k-1}=\sum_{j=1}^{\mu}\rho_j(k-\varepsilon_j)-\sum_{j=1}^{\mu}\rho_j(k-1-\varepsilon_j)=\sum_{j=1}^{\mu}\rho_j$

(δ) $\theta_k-\theta_{k-1}=\rho_1(k-\varepsilon_1)-0=\rho_1(\varepsilon_1+1-\varepsilon_1)=\rho_1$

(ε) $\theta_k=0$, $\theta_{k-1}=0$ and thus $\theta_k-\theta_{k-1}=0$ □

Corollary (5.4) is similar in nature with Corollary (4.4) and thus it provides the basis for the computation of $I_c(F,G)$ from the numbers θ_k. We first note that $\sum_{j=1}^{\mu}\rho_j=\rho=\dim N_r\{sF-\hat{s}G\}$; for all $k\geq\varepsilon_\mu+1$, $\theta_k-\theta_{k-1}=\rho$ and thus the smallest integer $k=\varepsilon_\mu+1$ for which $\theta_k-\theta_{k-1}=\rho$ may be defined by the rank tests and shall be referred to as the <u>right index</u> of (F,G). From the above results, we may deduce the following information about the right singular pencil.

<u>Remark (5.7)</u>: The differences $\theta_k-\theta_{k-1}$ are non-decreasing and the following properties hold true:

(i) There is an integer τ such that for $\forall k\geq\tau$, $\theta_k-\theta_{k-1}=\rho$. The integer $\rho=\dim N_r\{sF-\hat{s}G\}$ and the smallest integer τ for which $\theta_k-\theta_{k-1}=\rho$ is $\tau=\varepsilon_\mu+1$, the right index of (F,G).

(ii) The smallest index τ' for which $\theta_{\tau'}-\theta_{\tau'-1}\neq 0$ $(\theta_o=0)$ is $\tau'=\varepsilon_1+1$, where $\theta_{\tau'}-\theta_{\tau'-1}=\rho_1$ is the multiplicity of smallest c.m.i. ε_1.

(iii) The difference $\theta_k-\theta_{k-1}$ defines the total number of c.m.i. with value less or equal to k-1.

The sequence $C_r=\{\theta_k: k=-1,0,1,2,\ldots$, where $\theta_{-1}=\theta_o=0$, $\theta_k=\dim N_r^k(F,G)$, $k\geq 1\}$ is defined as the <u>right-(F,G) sequence</u> of $sF-\hat{s}G$. The properties of this sequence are defined by the following result.

<u>Corollary (5.5)</u>: For every k=0,1,2,..., the sequence C_r is characterised by the property

$$\theta_k \leq \frac{\theta_{k+1}+\theta_{k-1}}{2}\ ,\ \theta_{-1}=\theta_o=0 \tag{5.23}$$

In particular, we have that:

(i) Strict inequality holds, if and only if k=ε, where ε is a c.m.i.

(ii) Equality holds, if and only if k is not the value of a c.m.i.

Proof

The proof of this result follows by using Corollary (5.4) and by considering all possible cases which may occur. Thus we have:

(1) $\varepsilon_1=0$: $\theta_o-\theta_{-1}=0<\theta_1-\theta_o=\rho_1$.

(2) $\varepsilon_t+1\le k-1<k<k+1<\varepsilon_{t+1}+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{t}\rho_j=\theta_k-\theta_{k-1}$.

(3) $\varepsilon_t+1\le k-1<k<k+1=\varepsilon_{t+1}+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{t+1}\rho_j>\sum_{j=1}^{t}\rho_j=\theta_k-\theta_{k-1}$.

(4) $\varepsilon_{t-1}+1\le k-1<k=\varepsilon_t+1<k+1<\varepsilon_{t+1}+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{t}\rho_j=\theta_k-\theta_{k-1}$.

(5) $\varepsilon_{t-1}+1\le k-1<k=\varepsilon_t+1<k+1=\varepsilon_{t+1}+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{t+1}\rho_j>\sum_{j=1}^{t}\rho_j=\theta_k-\theta_{k-1}$.

(6) $\varepsilon_\mu+1\le k-1<k<k+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{\mu}\mu_j=\theta_k-\theta_{k-1}$.

(7) $\varepsilon_{\mu-1}+1\le k-1<k=\varepsilon_\mu+1<k+1$: $\theta_{k+1}-\theta_k=\sum_{j=1}^{\mu}\rho_j=\theta_k-\theta_{k-1}$.

(8) $k-1<k<k+1<\varepsilon_1+1$: $\theta_{k+1}-\theta_k=0=\theta_k-\theta_{k-1}$.

(9) $k-1<k<k+1=\varepsilon_1+1$: $\theta_{k+1}-\theta_k=\rho_1>0=\theta_k-\theta_{k-1}$.

(10) $k-1<\varepsilon_1+1=k<k+1<\varepsilon_2+1$: $\theta_{k+1}-\theta_k=\rho_1=\theta_k-\theta_{k-1}$.

(11) $k-1<\varepsilon_1+1=k<k+1=\varepsilon_2+1$: $\theta_{k+1}-\theta_k=\rho_1+\rho_2>\rho_1=\theta_k-\theta_{k-1}$.

From the above cases, it is clear that $\theta_{k+1}-\theta_k>\theta_k-\theta_{k-1}$ if and only if $k=\varepsilon$, where ε is a c.m.i. and that $\theta_{k+1}-\theta_k=\theta_k-\theta_{k-1}$ if and only if k is not a c.m.i.

□

Given that $\theta_k-\theta_{k-1}\ge 0$ for all $k=0,1,2,\ldots$, the sequence C_r is non-decreasing. For all integers k which are not values of c.m.i., the elements θ_k satisfy the arithmetic progression relationship $\theta_k=(\theta_{k+1}+\theta_{k-1})/2$. For those values of k which coincide with a c.m.i. the arithmetic progression relationship is violated since then $\theta_k<(\theta_{k+1}+\theta_{k-1})/2$. The sequence C_r is thus partitioned by the values of the c.m.i., i.e. the integers $0\le\varepsilon_1<\varepsilon_2<\ldots<\varepsilon_\mu$. For all k in the range of values $(\varepsilon_j,\ldots,\varepsilon_{j+1})$, starting from $k=\varepsilon_j+1$ and ending with $k=\varepsilon_{j+1}-1$, the relationship $\theta_k=(\theta_{k+1}+\theta_{k-1})/2$ holds true; this relationship, however, cannot be continued in the range of values $(\ldots,\varepsilon_j-1,\varepsilon_j)$, or $(\varepsilon_{j+1},\varepsilon_{j+1}+1,\ldots)$, since for $k=\varepsilon_j$, $\theta_{\varepsilon_j}<(\theta_{\varepsilon_j-1}+\theta_{\varepsilon_j+1})/2$ and for $k=\varepsilon_{j+1}$, $\theta_{\varepsilon_{j+1}}<(\theta_{\varepsilon_{j+1}}-1+\theta_{\varepsilon_{j+1}}+1)/2$. The sequence C_r therefore satisfies the arithmetic progression property in the range of values $(\varepsilon_j,\ldots,\varepsilon_{j+1})$ of k, but violates the arithmetic progression property at the boundary values $\varepsilon_j,\varepsilon_{j+1}$. The number $\delta_{\varepsilon_j}=(\theta_{\varepsilon_j+1}-\theta_{\varepsilon_j})-(\theta_{\varepsilon_j}-\theta_{\varepsilon_j-1})=\theta_{\varepsilon_j+1}+\theta_{\varepsilon_j-1}-2\theta_{\varepsilon_j}$ is a measure of deviation from the arithmetic progression type property at $k=\varepsilon_j$; the value $k=\varepsilon_j$ will be called a <u>singular point</u> of C_r and the number δ_{ε_j} will be called the <u>gap</u> of the sequence at $k=\varepsilon_j$. The sequence C_r will be regerred to as the <u>right Piecewise Arithmetic Progression sequence</u> (RPAPS) of (F,G) and it is clearly of similar nature to the PAPS of (F,G) at $s=\alpha$ defined on a regular pencil in Chapter (4). We may now state the following result relating the properties of C_r to the set $I_c(F,G)$ of the right singular pencil $sF-\hat{s}G$.

<u>Proposition (5.6)</u>: Let C_r be the RPAPS of (F,G). Then,

(i) An index $k=\varepsilon$ is a singular point of the sequence C_r, if and only if ε is the value of a c.m.i. of $sF-\hat{s}G$. Then, $\theta_\varepsilon<(\theta_{\varepsilon+1}+\theta_{\varepsilon-1})/2$.

(ii) If $k=\varepsilon$ is a singular point, then the gap $\delta_\varepsilon=\theta_{\varepsilon+1}+\theta_{\varepsilon-1}-2\theta_\varepsilon$ is equal to the multiplicity ρ of the c.m.i. with value ε.

<u>Proof</u>

Part (i) follows immediately by Corollary (5.5). If $k=\varepsilon_t$ is the value of a c.m.i., then $\theta_{k+1}-\theta_k=\sum_{j=1}^{t}\rho_j$ and $\theta_k-\theta_{k-1}=\sum_{j=1}^{t-1}\rho_j$ (by Corollary (5.4)); thus $\delta_k=\sum_{j=1}^{t}\rho_j-\sum_{j=1}^{t-1}\rho_j=\rho_j$. □

By finding the singular points of C_r and the corresponding gaps, the set of c.m.i. $I_c(F,G)$ is thus defined. The analysis presented so far leads to the following procedure for the determination of $I_c(F,G)$.

Piecewise Arithmetic Progression sequence diagram (PAPSD): Compute the numbers $\theta_{-1},\theta_o,\theta_1,\theta_2,\ldots,\theta_{\tilde{\sigma}+1}$, where $\theta_{-1}=\theta_o=0$, and $\tilde{\sigma}=m+1$, if $m<n$ and $\tilde{\sigma}=n$, if $m\geq n$. Compute then the gaps of RPAPS C_r, i.e.

$$\delta_k = \theta_{k+1}+\theta_{k-1}-2\theta_k, \quad k=0,1,2,\ldots,\tilde{\sigma}$$

and form a table of the following type: For every index k there is a value $\delta_k\geq 0$. If $\delta_k=0$, a dot is placed below δ_k and if $\delta_k>0$, then we create a column with asterisks below δ_k, with the number of asterisks being equal to the value of δ_k. This procredure is illustrated by the following diagram:

$$\begin{array}{lccccccccccc}
\text{index:} & 0, & 1, & 2, & \ldots, & k-1, & k, & k+1, & \ldots, & \tilde{\sigma}, & \tilde{\sigma}+1 \\
\text{gap:} & \delta_o, & \delta_1, & \delta_2, & \ldots, & \delta_{k-1}, & \delta_k, & \delta_{k+1}, & \ldots, & \delta_{\tilde{\sigma}}, & \delta_{\tilde{\sigma}+1} \\
 & * & \cdot & * & \ldots & \cdot & * & & \ldots & \cdot & \cdot \\
 & * & & * & & & * & & & & \\
 & & & * & & & \vdots & & & & \\
 & & & & & & * & & & &
\end{array}$$

Figure (5.1)

The indices characterised by dots do not correspond to values of c.m.i., whereas those characterised by asterisks define values of c.m.i. The number of asterisks in a column gives the multiplicity of the c.m.i., whose value is the corresponding index. Thus, for instance, in the above diagram we have c.m.i. with values, 0,2,k. □

The above procedure will be illustrated later on by an example. An alternative procedure for the computation of $I_c(F,G)$ is discussed next. This technique is similar to that presented for the computation of the Segre characteristic at s=α of a regular pencil and which has been based on the notion of Weyr characteristic and Ferrer's type diagram. For the sequence C_r we first define the following induced sequence:

$$W_r = \{\gamma_k : \gamma_k=\theta_k-\theta_{k-1}, \ k=0,1,2,\ldots\} \qquad (5.24)$$

The sequence $\mathcal{W}_r$ will be referred to as the <u>right $\mathcal{W}$eyr sequence</u> of (F,G) and its properties are characterised by the following result.

<u>Proposition (5.7)</u>: Let $\mathcal{W}_r$ be the right $\mathcal{W}$eyr sequence of (F,G). Then,

(i) $\gamma_k \leq \gamma_{k+1}$ for all k=0,1,2,... There always exists an integer τ, $\tau \leq \tilde{\sigma}$ ($\tilde{\sigma}=m+1$, if $m<n$, and $\tilde{\sigma}=n$, if $m \geq n$), such that for $\forall k \geq \tau$, $\gamma_k = \gamma_{k+1}$.

(ii) The strict inequality $\gamma_k < \gamma_{k+1}$ holds true, if and only if $k=\varepsilon$, where ε is the value of a c.m.i. The multiplicity ρ of ε is then defined by $\rho = \gamma_{k+1} - \gamma_k$.

□

This result is an alternative statement of the results established by Corollary (5.5) and Proposition (5.6). Proposition (5.7) suggests the following alternative procedure for the computation of $\mathcal{I}_c$(F,G).

<u>Weyr sequence diagram (WSD)</u>: Let $\tilde{\sigma}$ be the integer defined by $\tilde{\sigma}=m+1$, if $m<n$, and $\tilde{\sigma}=n$, if $m \geq n$. For every $\gamma_k \in \mathcal{W}_r$, $k=0,1,2,\ldots,\tilde{\sigma}$ we create a row in the following way: If $\gamma_k=0$, the row is filled in with dots and if $\gamma_k>0$, the row is filled with γ_k asterisks. The table created is parametrised by k and has the following general shape:

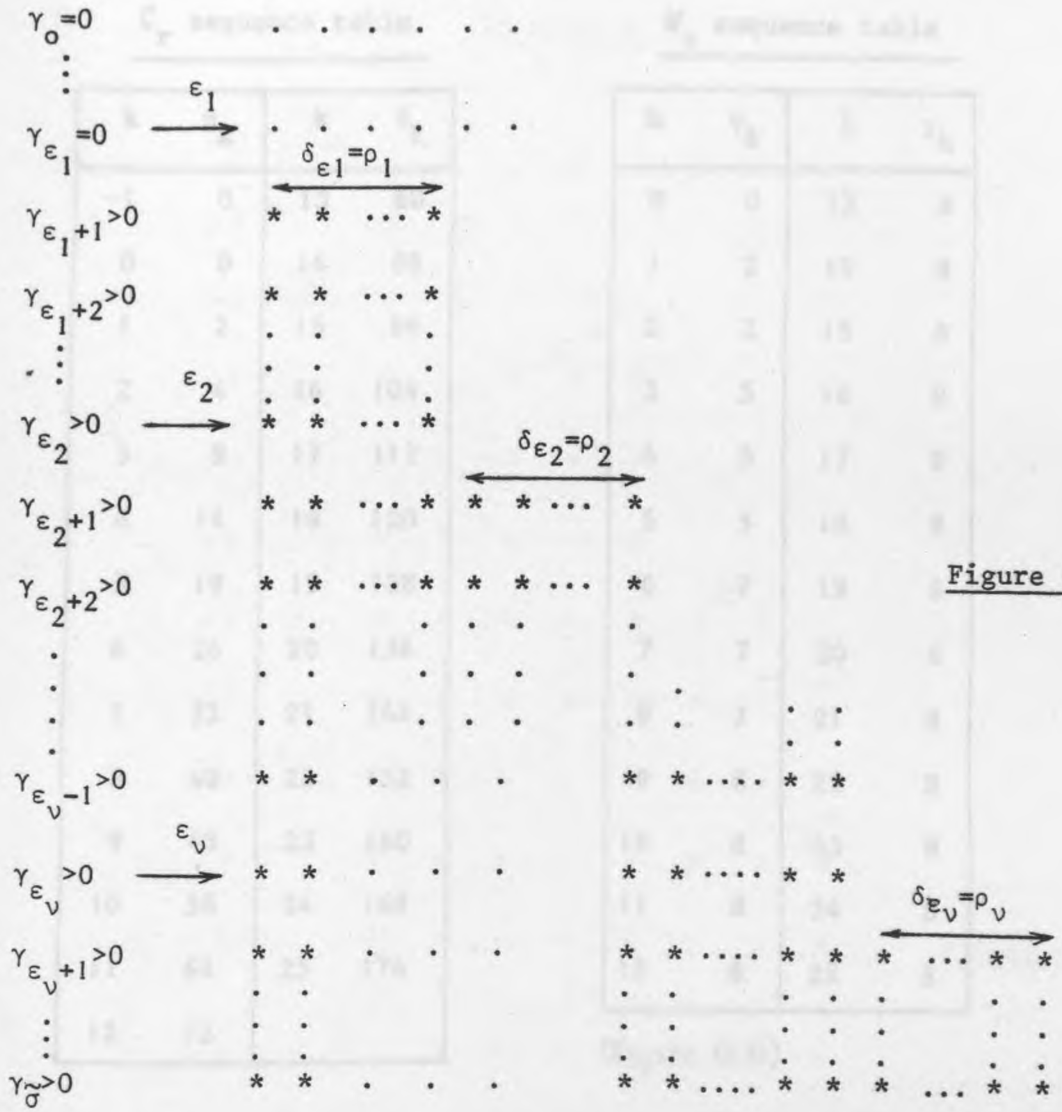

Figure (5.2)

From the above staircase diagram, the values of c.m.i. are computed as those integers associated with the different steps; the multiplicities of the corresponding c.m.i. are defined by the width (gap) $\delta_\varepsilon=\gamma_{\varepsilon+1}-\gamma_\varepsilon$ of the step. □

The following example illustrates the two procedures presented above.

Example (5.1): Let $sF-\hat{s}G$ be a 24×32 singular pencil and let the sequences $\mathcal{C}_r$ and $\mathcal{W}_r$ be given for $k=-1,0,1,\ldots,25$ $(\tilde{\sigma}=25)$ in the following tables:

C_r sequence table

k	θ_k	k	θ_k
-1	0	13	80
0	0	14	88
1	2	15	96
2	4	16	104
3	9	17	112
4	14	18	120
5	19	19	128
6	26	20	136
7	33	21	144
8	40	22	152
9	48	23	160
10	56	24	168
11	64	25	176
12	72		

W_r sequence table

k	γ_k	k	γ_k
0	0	13	8
1	2	14	8
2	2	15	8
3	5	16	8
4	5	17	8
5	5	18	8
6	7	19	8
7	7	20	8
8	7	21	8
9	8	22	8
10	8	23	8
11	8	24	8
12	8	25	8

(Figure (5.3))

From the above tables we may now form the RPAPS diagram and the WS diagram ; these two diagrams are shown in Figure (5.3) and Figure (5.4) respectively. The set of c.m.i. $I_c(F,G)$ is then given by

$$I_c(F,G) = \{(\varepsilon_1=0,\rho_1=2),(\varepsilon_2=2,\rho_2=3),(\varepsilon_3=5,\rho_3=2),(\varepsilon_4=8,\rho_4=1)\}$$

Note that since $\sum_{j=1}^{4} \rho_j(\varepsilon_j+1)=32$, the pencil $sF-\hat{s}G$ is entirely singular. □

RPAPS Diagram

```
index:  0  1  2  3  4  5  6  7  8  9 .... 24

gap:    2  0  3  0  0  2  0  0  1  0 .... 0
        *  .  *  .  .  *  .  .  *  .  ...  .
        *     *        *
              *
```

Figure (5.3)

WS Diagram

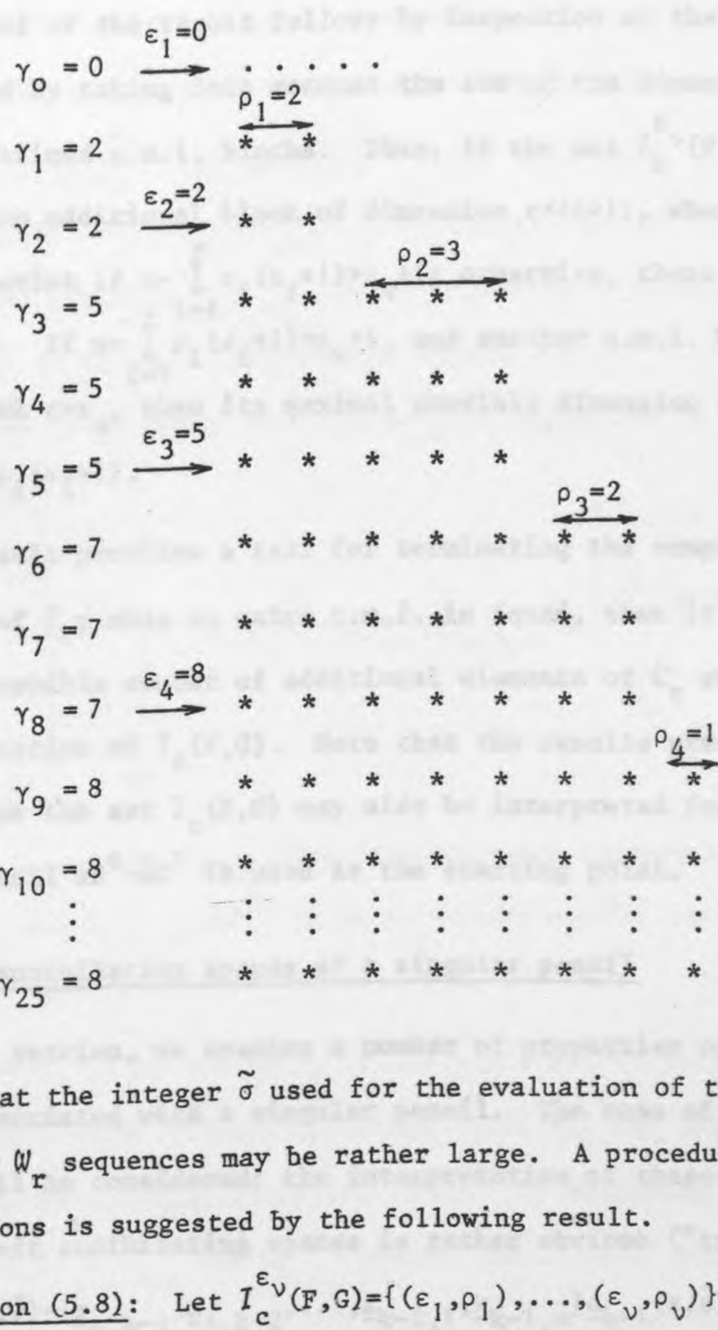

Figure (5.4)

Note that the integer $\tilde{\sigma}$ used for the evaluation of the first elements of C_r, or W_r sequences may be rather large. A procedure involving less computations is suggested by the following result.

Proposition (5.8): Let $I_c^{\varepsilon_\nu}(F,G)=\{(\varepsilon_1,\rho_1),\dots,(\varepsilon_\nu,\rho_\nu)\}\subseteq I_c(F,G)$ be the subset computed from the elements $\{\theta_{-1},\theta_o,\theta_1,\dots,\theta_{\varepsilon_\nu},\theta_{\varepsilon_\nu+1}\}$ of the C_r sequence of the m×n singular pencil $sF-\hat{s}G$. Then,

(i) $I_c^{\varepsilon_\nu}(F,G)=I_c(F,G)$ if $\pi_{\varepsilon_\nu}=n-\sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)-(\varepsilon_\nu+1)\leq 0$.

(ii) If $\pi_{\varepsilon_\nu}>0$, then at most the elements $\{\theta_{\varepsilon_\nu+2},\dots,\theta_{k+1}\}$ are needed for determining the $I_c(F,G)$ set, where $k=n-\sum_{i=1}\rho_i(\varepsilon_i+1)-1>\varepsilon_\nu+1$.

Proof

The proof of the result follows by inspection of the dimension of the pencil and by taking into account the sum of the dimensions of the already defined c.m.i. blocks. Thus, if the set $\mathcal{I}_c^{\varepsilon_\nu}(F,G)$ has been defined, an additional block of dimension $\varepsilon\times(\varepsilon+1)$, where $\varepsilon>\varepsilon_\nu$, may possibly exist if $n-\sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)>\varepsilon_\nu+1$; otherwise, there is no c.m.i. block with $\varepsilon>\varepsilon_\nu$. If $n-\sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)>\varepsilon_\nu+1$, and another c.m.i. block $\varepsilon\times(\varepsilon+1)$ exists with $\varepsilon>\varepsilon_\nu$, then its maximal possible dimension is k (k+1), where $k+1=n-\sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)$. □

This result provides a test for terminating the computation of the elements of $\mathcal{C}_r$; when an extra c.m.i. is found, then it also indicates the maximal possible number of additional elements of $\mathcal{C}_r$ which are needed for the computation of $\mathcal{I}_c(F,G)$. Note that the results presented in this section for the set $\mathcal{I}_c(F,G)$ may also be interpreted for the set $\mathcal{I}_r(F,G)$ if the pencil $sF^t-\hat{s}G^t$ is used as the starting point.

5.5 The annihilating spaces of a singular pencil

In this section, we examine a number of properties of the annihilating spaces associated with a singular pencil. The case of right annihilating spaces will be considered; the interpretation of these results to the case of left annihilating spaces is rather obvious ("transposed duality").

Let $\underline{x}(s,\hat{s})=[\underline{x}_{o,k-1},\underline{x}_{1,k-2},\ldots,\underline{x}_{k-2,1},\underline{x}_{k-1,o}]\underline{e}_{k-1}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ be a k-th right annihilating polynomial, generated by the k-th right annihilating vector $\underline{x}_k=[\underline{x}_{k-1,o}^t,\ldots,\underline{x}_{o,k-1}^t]^t$. The associated supporting subspace $\mathcal{X}(\underline{x}_k)=\text{span}\{\underline{x}_{k-1,o},\ldots,\underline{x}_{o,k-1}\}$ has been defined as the associated k-th right annihilating space of $\underline{x}_k$, and shall be denoted by $\mathcal{R}(\underline{x}_k)$.

The properties of $\mathcal{R}(\underline{x}_k)$ stem from the fact that $\underline{x}_k\in\mathcal{N}_r^k(F,G)$; the analysis in the previous sections thus provides the rools for the study of the properties of $\mathcal{R}(\underline{x}_k)$ spaces. Of special interest, is the parametrisation of the basis matrices X_{k-1}. This analysis leads to the definition of the maximal right annihilating space of $sF-\hat{s}G$ and to the introduction of a minimal dimension direct sum decomposition for this subspace.

Proposition (5.9): Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil and let $\mathcal{I}_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\le\ldots\le\varepsilon_\rho\}$be the set of c.m.i. Let us also assume that $\underline{x}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ be a general k-th right annihilating polynomial of $sF-\hat{s}G$. There always exists a $Q\in\mathbb{R}^{n\times n}$, $|Q|\neq 0$, such that the general family of basis matrices X_{k-1} is expressed by

$$X_{k-1}=Q\begin{bmatrix} D_{k,o,g} \\ \hline D_{k,\varepsilon_{g+1}} \\ \hline \vdots \\ \hline D_{k,\varepsilon_\nu} \\ \hline 0 \end{bmatrix},\ \text{if } \varepsilon_\nu+1\le k<\varepsilon_{\nu+1}+1,\ X_{k-1}=Q\begin{bmatrix} D_{k,o,g} \\ \hline D_{k,\varepsilon_{g+1}} \\ \hline \vdots \\ \hline D_{k,\varepsilon_\rho} \\ \hline 0 \end{bmatrix},\ \text{if } k\ge\varepsilon_\rho+1 \tag{5.25}$$

where D_{k,ε_j} are the (k,ε_j)-Toeplitz matrices corresponding to nonzero c.m.i. ε_j and $D_{k,o,g}\in\mathbb{R}^{g\times k}$ is an arbitrary matrix associated with the g zero c.m.i. □

This result is an extension of the basis matrix parametrisation result for entirely right singular matrices [Kar. -2] and its proof readily follows by Corollaries (5.1) and (5.3). By assigning arbitrarily the parameters in $D_{k,o,g}$ and D_{k,ε_j} blocks, families of supporting subspaces $\mathcal{R}(\underline{x}_k)$ are defined. The properties of supporting subspaces $\mathcal{R}(\underline{x}_k)$ are examined next.

Let $\underline{x}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ be a right annihilating polynomial vector. $\underline{x}(s,\hat{s})$ will be called *prime*, if $\text{rank}_{\mathbb{R}}\{X_{k-1}\}=k$; then $\mathcal{R}(\underline{x}_k)=\text{col-span}\{X_{k-1}\}$ has dimension k. The polynomial vector $\underline{x}(s,\hat{s})$ will be called *non-prime* if $\text{rank}\{X_{k-1}\}<k$. For non-prime annihilating vectors we have:

Lemma (5.3): Let $\underline{x}(s,\hat{s})=[\underline{x}_o,\underline{x}_1,\dots,\underline{x}_k]\underline{e}_k(s,\hat{s})=X_k\underline{e}_k(s,\hat{s})$ be a non-prime annihilating polynomial and let ν be the index for which the vector chain $\{\underline{x}_o,\underline{x}_1,\dots,\underline{x}_{\nu-1}\}$ is independent but $\{\underline{x}_o,\underline{x}_1,\dots,\underline{x}_{\nu-1},\underline{x}_\nu\}$ is dependent. There exists a prime right annihilating polynomial vector $\underline{x}^*(s,\hat{s})=[\underline{x}^*_o,\underline{x}^*_1,\dots,\underline{x}^*_{\nu-1}]\underline{e}_{\nu-1}(s,\hat{s})=X^*_{\nu-1}\underline{e}_{\nu-1}(s,\hat{s})$ for which $\text{span}\{\underline{x}^*_o,\dots,\underline{x}^*_{\nu-1}\}=\text{span}\{\underline{x}_o,\underline{x}_1,\dots,\underline{x}_{\nu-1}\}$.

Proof

Since $\underline{x}(s,\hat{s})\in N_r\{sF-\hat{s}G\}$, we have that

$$G\underline{x}_o=0,G\underline{x}_1=F\underline{x}_o,G\underline{x}_2=F\underline{x}_1,\dots,G\underline{x}_k=F\underline{x}_{k-1},F\underline{x}_k=0 \tag{5.26}$$

We may write $\underline{x}_\nu=\alpha_1\underline{x}_{\nu-1}+\alpha_2\underline{x}_{\nu-2}+\dots+\alpha_\nu\underline{x}_\nu$ and thus by (5.26) we have

$$G\underline{x}_\nu=\alpha_1 G\underline{x}_{\nu-1}+\alpha_2 G\underline{x}_{\nu-2}+\dots+\alpha_{\nu-1}G\underline{x}_1+\alpha_\nu\underbrace{G\underline{x}_o}_{\overset{\shortparallel}{0}}$$

or

$$F\underline{x}_{\nu-1}=\alpha_1 F\underline{x}_{\nu-2}+\alpha_2 F\underline{x}_{\nu-3}+\dots+\alpha_{\nu-1}F\underline{x}_o$$

Then, it is clear that

$$F\underline{x}^*_{\nu-1}=0,\text{ where }\underline{x}^*_{\nu-1}=\underline{x}_{\nu-1}-\alpha_1\underline{x}_{\nu-2}-\alpha_2\underline{x}_{\nu-3}-\dots-\alpha_{\nu-1}\underline{x}_o \tag{5.27a}$$

By (5.27a) and (5.26) it follows that

$$G\underline{x}^*_{\nu-1}=G\underline{x}_{\nu-1}-\alpha_1 G\underline{x}_{\nu-2}-\dots-\alpha_{\nu-1}\underbrace{G\underline{x}_o}_{\overset{\shortparallel}{0}}$$

$$=F\{\underline{x}_{\nu-2}-\alpha_1\underline{x}_{\nu-3}-\alpha_2\underline{x}_{\nu-4}-\dots-\alpha_{\nu-2}\underline{x}_o\}$$

and thus

$$G\underline{x}^*_{\nu-1}=F\underline{x}^*_{\nu-2},\text{ where }\underline{x}^*_{\nu-2}=\underline{x}_{\nu-2}-\alpha_1\underline{x}_{\nu-3}-\dots-\alpha_{\nu-2}\underline{x}_o \tag{5.27b}$$

We may continue the process and define the vectors

$$\overset{*}{\underline{x}}_{\nu-3}=\underline{x}_{\nu-3}-\alpha_1\underline{x}_{\nu-4}-\dots-\alpha_{\nu-3}\underline{x}_o,\dots,\overset{*}{\underline{x}}_1=\underline{x}_1-\alpha_1\underline{x}_o,\overset{*}{\underline{x}}_o=\underline{x}_o \qquad (5.27c)$$

and the vectors $\{\overset{*}{\underline{x}}_o,\overset{*}{\underline{x}}_1,\dots,\overset{*}{\underline{x}}_{\nu-3},\overset{*}{\underline{x}}_{\nu-2},\overset{*}{\underline{x}}_{\nu-1}\}$ satisfy the conditions

$$G\overset{*}{\underline{x}}_o=0,G\overset{*}{\underline{x}}_1=F\overset{*}{\underline{x}}_o,\dots,G\overset{*}{\underline{x}}_{\nu-1}=F\overset{*}{\underline{x}}_{\nu-2},F\overset{*}{\underline{x}}_{\nu-1}=0 \qquad (5.27d)$$

Thus the vector $\overset{*}{\underline{x}}(s,\hat{s})=[\overset{*}{\underline{x}}_o,\overset{*}{\underline{x}}_1,\dots,\overset{*}{\underline{x}}_{\nu-1}]\underline{e}_{\nu-1}(s,\hat{s})$ is a right annihilating vector of degree $\nu-1$. Note that

$$[\overset{*}{\underline{x}}_o,\overset{*}{\underline{x}}_1,\dots,\overset{*}{\underline{x}}_{\nu-1}]=[\underline{x}_o,\underline{x}_1,\dots,\underline{x}_{\nu-1}]\begin{bmatrix} 1 & -\alpha_1 & \dots & & -\alpha_{\nu-1} \\ 0 & 1 & -\alpha_1 & \cdots & -\alpha_{\nu-2} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 1 & -\alpha_1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \triangleq P \qquad (5.28$$

Since P has full rank, $\{\overset{*}{\underline{x}}_o,\dots,\overset{*}{\underline{x}}_{\nu-1}\}$ are linearly independent and $sp\{\overset{*}{\underline{x}}_o,\overset{*}{\underline{x}}_1,\dots,\overset{*}{\underline{x}}_{\nu-1}\}=sp\{\underline{x}_o,\underline{x}_1,\dots,\underline{x}_{\nu-1}\}$. □

The above Lemma is implicit in the reduction of a pencil to the Kronecker form [Gan. -1].

Proposition (5.10): Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil. There exists a minimal basis $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}),\ i\in\underset{\sim}{p}\}$ for $\mathcal{N}_r\{sF-\hat{s}G\}$ with the following properties:

(i) The vectors $\underline{x}_{\varepsilon_i}(s,\hat{s})=X_{\varepsilon_i}\cdot\underline{e}_{\varepsilon_i}(s,\hat{s})$ are prime for $\forall\ i\in\underset{\sim}{p}$.

(ii) The set of subspaces $\mathcal{R}(\mathcal{B})=\{\mathcal{R}_{\varepsilon_i}:\mathcal{R}_{\varepsilon_i}=sp\{X_{\varepsilon_i}\},\ i\in\underset{\sim}{p}\}$ is linearly independent.

Proof

Let $\underline{e}_{\varepsilon}(s,\hat{s})=[\hat{s}^{\varepsilon},\hat{s}^{\varepsilon-1}s,\dots,\hat{s}s^{\varepsilon-1},s^{\varepsilon}]^t$ be the ε-th order vector of a-polarity, when $\varepsilon>0$ is a nonzero c.m.i. and let $\underline{e}_{\varepsilon}(s,\hat{s})=1$, when $\varepsilon=0$. The matrix

$$E(s,\hat{s}) = \left[\begin{array}{ccccc} \underline{e}_{\varepsilon_1}(s,\hat{s}) & & & & \\ & \ddots & & \bigcirc & \\ & & \underline{e}_{\varepsilon_i}(s,\hat{s}) & & \\ \bigcirc & & & \ddots & \\ & & & & \underline{e}_{\varepsilon_p}(s,\hat{s}) \\ \hdashline & & \bigcirc & & \end{array}\right] \begin{array}{l} \uparrow \\ \sum_{i=1}^{p}(\varepsilon_i+1)=\sigma \\ \downarrow \ \in\mathbb{R}^{n\times\sigma}[s,\hat{s}] \end{array} \tag{5.29}$$

defines a minimal basis for $N_r\{sF_k-\hat{s}G_k\}$. If (R,Q) is a pair of strict equivalence transformations that reduce $sF-\hat{s}G \xrightarrow{(R,Q)} R(sF-\hat{s}G)Q-sF_k-\hat{s}G_k$, then the columns of the matrix

$$X(s,\hat{s})=QE(s,\hat{s})=[\underline{x}_{\varepsilon_1}(s,\hat{s}),\ldots,\underline{x}_{\varepsilon_i}(s,\hat{s}),\ldots,\underline{x}_{\varepsilon_p}(s,\hat{s})] \tag{5.30a}$$

are linearly independent over $\mathbb{R}(s,\hat{s})$ and since the degrees are ε_i, $i\in\underset{\sim}{p}$, they define a minimal basis for $N_r\{sF-\hat{s}G\}$. If we partition Q as

$$Q=[\underset{\substack{\longleftrightarrow\\ \varepsilon_1+1}}{X_{\varepsilon_1}};\ldots;\underset{\substack{\longleftrightarrow\\ \varepsilon_i+1}}{X_{\varepsilon_i}};\ldots;\underset{\substack{\longleftrightarrow\\ \varepsilon_p+1}}{X_{\varepsilon_p}};Q'] \tag{5.30b}$$

then $\text{rank}\{X_{\varepsilon_i}\}=\varepsilon_i+1$ and X_{ε_i} is a basis matrix for the minimal degree ε_i right annihilating vector $\underline{x}_{\varepsilon_i}(s,\hat{s})$; the set of supporting subspaces $\{R_{\varepsilon_i} : R_{\varepsilon_i}=\text{sp}\{X_{\varepsilon_i}\},\ i\in\underset{\sim}{p}\}$ is clearly linearly independent and since $\dim R_{\varepsilon_i}=\varepsilon_i+1$, the vectors $\underline{x}_{\varepsilon_i}(s,\hat{s})$ are prime. □

The basis matrix $E(s,\hat{s})\in\mathbb{R}^{n\times\sigma}[s,\hat{s}]$ will be referred to as the <u>right-structure matrix</u> of $sF-\hat{s}G$. A set of vectors $S(s,\hat{s})=\{\underline{x}_i(s,\hat{s}): \underline{x}_i(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}],\ i\in\underset{\sim}{r}\}$ with the vectors $\underline{x}_i(s,\hat{s})$ linearly independent over $\mathbb{R}(s,\hat{s})$ will be referred to as an <u>r-rank set</u> of $\mathbb{R}^n[s,\hat{s}]$. If the vectors of the r-rank set $S(s,\hat{s})$ are all prime, then $S(s,\hat{s})$ will be called an <u>r-rank prime set</u> of $\mathbb{R}^n[s,\hat{s}]$. With every vector $\underline{x}_i(s,\hat{s})= =X_{p_i}\underline{e}_{p_i}(s,\hat{s})$ of $S(s,\hat{s})$, we associate the supporting subspace $R_{p_i}=\text{sp}\{X_{p_i}\}$; the linear vector space $R_S=\sum_{i=1}^{r}R_{p_i}$ will be called the <u>characteristic space</u> of $S(s,\hat{s})$. If $S(s,\hat{s})$ is an r-rank prime set and the subspaces $\{R_{p_i},\ i\in\underset{\sim}{r}\}$ are linearly independent, then $S(s,\hat{s})$ will be called <u>complete</u>. With a rank r set $S(s,\hat{s})$ we may associate two modules; by setting $\hat{s}=1$, and $s=1$ respectively, we obtain the sets $S(s)\triangleq S(s,1)=\{\underline{x}_i(s): \underline{x}_i(s)=\underline{x}_i(s,1)\in\mathbb{R}^n[s], i\in\underset{\sim}{r}\}$ and $\hat{S}(\hat{s})\triangleq S(1,\hat{s})=\{\underline{x}_i(\hat{s}): \underline{x}_i(\hat{s})=\underline{x}_i(1,\hat{s})\in\mathbb{R}^n[\hat{s}], i\in\underset{\sim}{r}\}$. The sets $S(s)$, $\hat{S}(\hat{s})$ have the same characteristic space, which is the space associated with $S(s,\hat{s})$. The finitely generated $\mathbb{R}[s]$, $\mathbb{R}[\hat{s}]$ modules defined by the sets $S(s),\hat{S}(\hat{s})$ respectively shall be denoted by $M_S,M_{\hat{S}}$ correspondingly and shall be referred to as the associated modules of $S(s,\hat{s})$. An important property of $\mathcal{R}_S$ is defined below.

Lemma (5.4): Let $S(s,\hat{s})=\{\underline{x}_i(s,\hat{s}): \underline{x}_i(s,\hat{s})=X_{p_i}\underline{e}_{p_i}(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}], i\in\underset{\sim}{r}\}$ be an r-rank set, $\mathcal{R}_S$ be the corresponding characteristic space and let $M_S,M_{\hat{S}}$ be the $\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$-associated modules with the set $S(s,\hat{s})$. Then,

(i) $\mathcal{R}_S$ is an invariant of both M_S and $M_{\hat{S}}$ modules.

(ii) If $\mathcal{R}_{p_i}$ are the supporting subspaces of $\underline{x}_{p_i}(s,\hat{s})$, then

$$\dim\mathcal{R}_S\le\dim\mathcal{R}_{p_1}+\dots+\dim\mathcal{R}_{p_r}$$

and there is equality, if and only if the subspaces $\{\mathcal{R}_{p_i}, i\in\underset{\sim}{r}\}$ are linearly independent.

Proof

(i) Let $S'(s)=\{\underline{x}'_i(s), i\in\underset{\sim}{r}\}$ be another basis of M_S. Then,

$$\underline{x}'_j(s)=\sum_{i=1}^{r}\alpha_{ji}(s)\underline{x}_i(s)=X'_{q_j}\underline{e}_{q_j}(s),\ \forall j\in\underset{\sim}{r},\text{ where }\alpha_{ji}(s)\in\mathbb{R}[s] \tag{5.31a}$$

If $\mathcal{R}'_{q_j}$ is the supporting subspace of $\underline{x}'_j(s)$, then (5.31a) implies that $\mathcal{R}'_{q_j}\subseteq\mathcal{R}_S$ and thus $\sum_{i=1}^{r}\mathcal{R}'_{q_j}=\mathcal{R}_{S'}\subseteq\mathcal{R}_S$, where $\mathcal{R}_{S'}$ is the characteristic space of $S'(s)$. Similarly, since $S'(s)$ is a basis of M_S we have

$$\underline{x}_j(s)=\sum_{i=1}^{r}b_{ji}(s)\underline{x}'_i(s),\ \forall j\in\underset{\sim}{r},\text{ where }b_{ji}(s)\in\mathbb{R}[s] \tag{5.31b}$$

and thus $\mathcal{R}_{p_i}\subseteq\mathcal{R}_{S'}$; the latter implies that $\sum_{i=1}^{r}\mathcal{R}_{p_i}=\mathcal{R}_S\subseteq\mathcal{R}_{S'}$, and therefore $\mathcal{R}_S=\mathcal{R}_{S'}$. The proof for $M_{\hat{S}}$ is identical.

(ii) This part is a straightforward application of a standard result [God. -1]. □

The notion of the characteristic space introduced for an r-rank set $\mathcal{S}(s,\hat{s})$ is now specialised to the minimal basis sets of $\mathcal{N}_r\{sF-\hat{s}G\}$. We may state the following result.

Theorem (5.3): Let $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}): \underline{x}_{\varepsilon_i}(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}],\ i\in\underline{p}\}$ be a minimal basis for $\mathcal{N}_r\{sF-\hat{s}G\}$ and let $\mathcal{R}^*$ be the characteristic space of $\mathcal{B}(s,\hat{s})$. Then the following properties hold true:

(i) $\mathcal{R}^*$ is an invariant of the rational vector space $\mathcal{N}_r\{sF-\hat{s}G\}$ and $\dim\mathcal{R}^*=\sum_{i=1}^{p}(\varepsilon_i+1)$.

(ii) Every minimal basis $\mathcal{B}(s,\hat{s})$ is complete.

(iii) The characteristic space $\mathcal{R}^*$ is the maximal right annihilating space of $sF-\hat{s}G$. If $\mathcal{R}_{\varepsilon_i}$ is the supporting subspace of a minimal basis $\mathcal{B}(s,\hat{s})$, then $\mathcal{R}^*$ may be decomposed as

$$\mathcal{R}^*=\mathcal{R}_{\varepsilon_1}\oplus\mathcal{R}_{\varepsilon_2}\oplus\ldots\oplus\mathcal{R}_{\varepsilon_p} \tag{5.32}$$

Proof

(i) By Lemma (5.4), all minimal bases of $\mathcal{N}_r\{sF-G\}$, $\mathcal{N}_r\{F-\hat{s}G\}$ have the same characteristic space $\mathcal{R}^*$, which is the characteristic space of the $\mathbb{R}[s,\hat{s}]$ minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$. By Proposition (5.10), there exists a complete minimal basis for $\mathcal{N}_r\{sF-\hat{s}G\}$. If $\mathcal{R}_{\varepsilon_i}$ are the supporting subspaces of the $\underline{x}_{\varepsilon_i}(s,s)$ vectors of this basis, then $\mathcal{R}^*=\mathcal{R}_{\varepsilon_1}\oplus\ldots\oplus\mathcal{R}_{\varepsilon_p}$ and thus $\dim\mathcal{R}^*=\sum_{i=1}^{p}\dim\mathcal{R}_{\varepsilon_i}=\sum_{i=1}^{p}(\varepsilon_i+1)$.

(ii) Consider the complete minimal basis $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}): \underline{x}_{\varepsilon_i}(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}],\ i\in\underline{p}\}$ and another minimal basis $\mathcal{B}'(s,\hat{s})=\{\underline{x}'_{\varepsilon_i}(s,\hat{s}): \underline{x}'_{\varepsilon_i}(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}],\ i\in\underline{p}\}$. If $\mathcal{R}'_{\varepsilon_i}$ are the supporting subspaces associated with the minimal degree ε_i vectors $\underline{x}'_{\varepsilon_i}(s,\hat{s})$, then clearly

$\rho_i = \dim \mathcal{R}'_{\varepsilon_i} \le (\varepsilon_i+1), \forall\, i\in\underline{p}$; however, by the invariance of the characteristic space (part (i)) we have that $\mathcal{R}^* = \sum_{i=1}^{p} \mathcal{R}'_{\varepsilon_i}$ and thus

$$\dim \mathcal{R}^* = \sum_{i=1}^{p} (\varepsilon_i+1) \le \sum_{i=1}^{p} \dim \mathcal{R}'_{\varepsilon_i} = \sum_{i=1}^{p} \rho_i \tag{5.33a}$$

since $\rho_i \le (\varepsilon_i+1)$, it follows that

$$\sum_{i=1}^{p} \rho_i = \sum_{i=1}^{p} \dim \mathcal{R}'_{\varepsilon_i} \le \sum_{i=1}^{p} (\varepsilon_i+1) = \dim \mathcal{R}^* \tag{5.33b}$$

By (5.33a) and (5.33b) it follows that

$$\dim \mathcal{R}^* = \sum_{i=1}^{p} (\varepsilon_i+1) = \sum_{i=1}^{p} \dim \mathcal{R}'_{\varepsilon_i} \tag{5.33c}$$

By part (ii) of Lemma (5.4) it follows that the subspaces $\{\mathcal{R}'_{\varepsilon_i}, i\in\underline{p}\}$ are linearly independent; furthermore, since $\sum_{i=1}^{p} (\varepsilon_i+1) = \sum_{i=1}^{p} \rho_i$ and $\rho_i \le (\varepsilon_i+1)$ we have that $\rho_i = \varepsilon_i+1, \forall\, i\in\underline{p}$ and thus the basis $\mathcal{B}'(s,\hat{s})$ is complete.

(iii) If $\mathcal{B}(s,\hat{s})$ is any minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$ and $\underline{x}(s,\hat{s})$ is an arbitrary right annihilating polynomial vector, then $\underline{x}(s,1)=\underline{x}(s)$ (or $\underline{x}(1,\hat{s})=\underline{\hat{x}}(\hat{s})$) may be expressed as

$$\underline{x}(s) = X_q \underline{e}_q(s) = \sum_{i=1}^{p} \alpha_i(s) \underline{x}_{\varepsilon_i}(s), \; \alpha_i(s) \in \mathbb{R}[s] \tag{5.34}$$

from which $\mathrm{sp}\{X_q\} \subseteq \mathcal{R}^*$. This proves that $\mathcal{R}^*$ is the maximal right annihilating space of $sF-\hat{s}G$. The direct sum decomposition of $\mathcal{R}^*$ follows from the fact that any minimal basis is complete. □

This result establishes that the characteristic subspace $\mathcal{R}^*$ associated with a minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$ is an invariant of the rational vector space $\mathcal{N}_r\{sF-\hat{s}G\}$; furthermore, any minimal basis is complete and thus defines a direct sum decomposition for $\mathcal{R}^*$, which however is not uniquely defined. The set of supporting subspaces $\{\mathcal{R}_{\mathcal{B}}\} = \{\mathcal{R}_{\varepsilon_i}, i\in\underline{p}\}$ associated with a minimal basis $\mathcal{B}(s,\hat{s})$ is linearly independent and their dimensions are minimal; such a set $\{\mathcal{R}_{\mathcal{B}}\}$ will be referred to as the minimal right annihilating set of $\mathcal{B}(s,\hat{s})$. The characteristic space $\mathcal{R}^*$ will be referred to as the <u>maximal right annihilating space</u> of $sF-\hat{s}G$. $\mathcal{R}^*$ may be computed without resorting to the computation of a minimal basis; furthermore, minimal bases may be computed without the use of algebraic tools, but by using the properties of the k-th right characteristic spaces of (F,G). Some further properties of the minimal bases are considered first.

<u>Corollary (5.6)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil, $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}): \underline{x}_{\varepsilon_i}(s,\hat{s})=[\underline{x}_o^i,\underline{x}_1^i,\dots,\underline{x}_{\varepsilon_i}^i]\underline{e}_{\varepsilon_i}(s,\hat{s})=X_{\varepsilon_i}\underline{e}_{\varepsilon_i}(s,\hat{s}),\ i\in\underline{p}\}$ be a minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$ and let $\mathcal{R}_\ell=\mathrm{sp}\{\underline{x}_o^1,\dots,\underline{x}_o^p\}$ and $\mathcal{R}_h=\mathrm{sp}\{\underline{x}_{\varepsilon_1}^1,\dots,\underline{x}_{\varepsilon_p}^p\}$, be the subspaces of $\mathcal{R}^*$. Then,

(i) The subspaces $\mathcal{R}_\ell,\mathcal{R}_h$ are invariants of the rational vector space $\mathcal{N}_r\{sF-\hat{s}G\}$.

(ii) $\mathcal{N}_r\{G\}\cap\mathcal{R}^*=\mathcal{R}_\ell$ and $\mathcal{N}_r\{F\}\cap\mathcal{R}^*=\mathcal{R}_h$.

<u>Proof</u>

(i) Let $\underline{x}(s,\hat{s})\in\mathbb{R}^n[s,\hat{s}]\in\mathcal{N}_r\{sF-\hat{s}G\}$ be a homogeneous polynomial vector of homogeneous degree k. For $\hat{s}=1$, $\underline{x}(s)=\underline{x}(s,1)$ may be expressed in terms of the vectors of the minimal basis $\mathcal{B}(s,1)$ as

$$\underline{x}(s)=\sum_{j:\varepsilon_j\le k} c_{\varepsilon_j}(s)\underline{x}_{\varepsilon_j}(s) \tag{5.35}$$

where $c_{\varepsilon_j}(s)\in\mathbb{R}[s]$ and $\deg\{c_{\varepsilon_j}(s)\}=k-\varepsilon_j$. If $\underline{x}(s)=\underline{x}_o+\underline{x}_1s+\dots+\underline{x}_ks^k$, then (5.35) implies that $\underline{x}_o\in\mathcal{R}_\ell$. If $\mathcal{B}'(s,\hat{s})=\{\underline{x}'_{\varepsilon_i}(s,\hat{s}),\ i\in\underline{p}\}$ is another minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$, then every vector $\underline{x}'_{\varepsilon_j}(s)$ may be expressed as in (5.35) and thus $\underline{x}_o'^j\in\mathcal{R}_\ell$; this clearly implies that $\mathcal{R}'_\ell=\mathrm{sp}\{\underline{x}_o'^1,\dots,\underline{x}_o'^p\}\subseteq\mathcal{R}_\ell$. However, every vector of $\mathcal{B}(s,1)$ may be expressed in terms of the basis $\mathcal{B}'(s,1)$ as in (5.35) and thus $\mathcal{R}_\ell\subseteq\mathcal{R}'_\ell$. By the conditions $\mathcal{R}'_\ell\subseteq\mathcal{R}_\ell$ and $\mathcal{R}_\ell\subseteq\mathcal{R}'_\ell$ it follows $\mathcal{R}_\ell=\mathcal{R}'_\ell$, which proves the invariance. The invariance of $\mathcal{R}_h$ is proved along similar lines (i.e. set s=1 and consider the vector $\underline{x}(\hat{s})=\underline{x}(1,\hat{s})$ etc.).

(ii) Clearly $\mathcal{N}_r\{G\}\cap\mathcal{R}^*=\tilde{\mathcal{R}}\neq\{0\}$ and $\mathcal{R}_\ell\subseteq\tilde{\mathcal{R}}$. To prove that $\mathcal{R}_\ell=\tilde{\mathcal{R}}$ we have to show that $N_r\{G\}\cap R^* \subseteq R_\ell$. In order to establish the latter condition, it is sufficient to show that the general vector $\underline{z}_o \in N_r\{G\}\cap R^*$ is the constant coefficient vector of some $\underline{z}(s,\hat{s}) \in N_r\{sF-\hat{s}G\}$ homogeneous polynomial vector; if this is shown then from the proof of part (i) it follows that $\underline{z}_o \in R_\ell$.

Let $\underline{z}_o \in N_r\{G\}\cap R^*$ and let $\tau=\{\dots;\underline{x}_o^i,\underline{x}_1^i,\dots,\underline{x}_{\varepsilon_i}^i;\dots\}$ be the basis for R^* defined by the coefficient of minimal basis vectors $\underline{x}_{\varepsilon_i}(s,\hat{s})$ of $B(s,\hat{s})$. Then,

$$\underline{z}_o = \sum_{j=0}^{\varepsilon_1} \alpha_j^1 \underline{x}_j^1 + \dots + \sum_{j=0}^{\varepsilon_p} \alpha_j^p \underline{x}_j^p \quad \text{and} \quad G\underline{z}_o = \underline{0} \tag{5.36a}$$

Note that

$$F\underline{z}_o = \sum_{j=0}^{\varepsilon_1-1} \alpha_j^1 F\underline{x}_j^1 + \dots + \sum_{j=0}^{\varepsilon_p-1} \alpha_j^p F\underline{x}_j^p = \sum_{j=1}^{\varepsilon_1} \alpha_{j-1}^1 G\underline{x}_j^1 + \dots + \sum_{j=1}^{\varepsilon_p} \alpha_{j-1}^p G\underline{x}_j^p \tag{5.36b}$$

since the vectors $\{\underline{x}_o^i,\dots,\underline{x}_{\varepsilon_i}^i\}$ satisfy conditions (5.4a). If we define the vector $\underline{z}_1 = \sum_{j=1}^{\varepsilon_1} \alpha_{j-1}^1 \underline{x}_j^1 + \dots + \sum_{j=1}^{\varepsilon_p} \alpha_{j-1}^p \underline{x}_j^p$, then $F\underline{z}_o = G\underline{z}_1$. We may continue the process and define vectors

$$\underline{z}_k = \sum_{j=k}^{\varepsilon_1} \alpha_{j-k}^1 \underline{x}_j^1 + \dots + \sum_{j=k}^{\varepsilon_p} \alpha_{j-k}^p \underline{x}_j^p \tag{5.36c}$$

where $\alpha_{j-k}^i=0$ if $k>\varepsilon_i$. The vectors $\underline{z}_k$ satisfy the conditions

$$G\underline{z}_o=0, G\underline{z}_1=F\underline{z}_o, \dots, G\underline{z}_{\varepsilon_p}=F\underline{z}_{\varepsilon_p-1}, 0=F\underline{z}_{\varepsilon_p} \tag{5.36d}$$

and thus there exists a homogeneous polynomial vector $\underline{z}(s,\hat{s})=\underline{z}_o\hat{s}^{\varepsilon_p} + \underline{z}_1\hat{s}^{\varepsilon_p-1}s + \dots + \underline{z}_{\varepsilon_p}s^{\varepsilon_p} \in N_r\{sF-\hat{s}G\}$ for every $\underline{z}_o \in N_r\{G\}\cap R^*$. The proof of the second statement of part (ii) is similar. □

<u>Remark (5.8)</u>: Let $\underline{x}(s,\hat{s})=\underline{x}_o\hat{s}^k+\underline{x}_1\hat{s}^{k-1}s+\dots+\underline{x}_{k-1}\hat{s}s^{k-1}+\underline{x}_k s^k$ be a general k-th right annihilating polynomial vector of $sF-\hat{s}G$. Then, $\underline{x}_o \in R_\ell$ and $\underline{x}_k \in R_h$.

<u>Corollary (5.6)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil, $B(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}): \underline{x}_{\varepsilon_i}(s,\hat{s})=[\underline{x}_o^i,\underline{x}_1^i,\dots,\underline{x}_{\varepsilon_i}^i]\underline{e}_{\varepsilon_i}(s,\hat{s})=X_{\varepsilon_i}\underline{e}_{\varepsilon_i}(s,\hat{s}),\ i\in\underline{p}\}$ be a homogeneous minimal basis of $N_r\{sF-\hat{s}G\}$ and let $T_{\varepsilon_i}=\text{col-sp}\{GX_{\varepsilon_i}\}$ for $\forall i\in\underline{p}$.

Then:

(i) The set of vectors $\{\underline{z}_j^i : \underline{z}_j^i = G\underline{x}_j^i \in \mathbb{R}^m,\ j \in \underset{\sim}{\varepsilon}_i\}$ form a basis for the subspace $\mathcal{T}_{\varepsilon_i}$, for ∀ $i \in \underset{\sim}{p}$, for which $\varepsilon_i > 0$.

(ii) The set of subspaces $\{\mathcal{T}_{\varepsilon_i} :$ for ∀ $\varepsilon_i > 0\}$ is linearly independent.

(iii) If $\{0 < \varepsilon_{g+1} \leq \ldots \leq \varepsilon_p\}$ is the set of nonzero c.m.i., then the subspace

$$\mathcal{T}^* = \mathcal{T}_{\varepsilon_{g+1}} \oplus \ldots \oplus \mathcal{T}_{\varepsilon_p} \tag{5.37}$$

is an invariant of $N_r\{sF - \hat{s}G\}$.

Proof

(i) Clearly, $G\underline{x}_o^i = 0$ and the vectors $\{\underline{z}_j^i = G\underline{x}_j^i,\ j \in \varepsilon_i\}$ are nonzero. Assume that the vectors $\{\underline{z}_j^i : j \in \underset{\sim}{\varepsilon}_i\}$ are dependent and let $G\underline{x}_h$ $(h \geq 1)$ be the first vector in $\{\underline{z}_1^i, \ldots, \underline{z}_h^i, \ldots\}$ that is linearly dependent on the preceding ones and let

$$G\underline{x}_h^i = \alpha_1 G\underline{x}_{h-1}^i + \alpha_2 G\underline{x}_{h-2}^i + \ldots + \alpha_{h-1} G\underline{x}_1^i$$

Following identical steps to those suggested in [Gant. -1] (Vol.2, pp32) it follows that vectors $\underline{\tilde{x}}_{h-k}^i = \underline{x}_{h-k}^i - \alpha_1 \underline{x}_{h-k-1}^i - \ldots - \alpha_{h-k} \underline{x}_o^i$, $k = 1, \ldots, h$ may be defined for which conditions (5.4a) are satisfied. Clearly, the vector $\underline{\tilde{x}}_{h-1}(s,s) = \tilde{X}_{h-1} \underline{e}_{h-1}(s,s) \in N_r\{sF - \hat{s}G\}$, $\mathrm{sp}\{\tilde{X}_{h-1}\} \subset R_{\varepsilon_i}$ and $\deg\{\underline{\tilde{x}}_{h-1}(s,\hat{s})\} = h-1 < \varepsilon_i$. If $h-1 < \min\{\varepsilon_j, j \in \underset{\sim}{p}\}$, then the existence of $\underline{\tilde{x}}_{h-1}(s,\hat{s})$ violates the minimality assumption for the $I_c(F,G)$ set of indices. If $h-1 \geq \min\{\varepsilon_j, j \in \underset{\sim}{p}\}$ and ε_ν is the maximal minimal index for which $\varepsilon_1 > h-1 \geq \varepsilon_\nu$, then the polynomial vector $\underline{\tilde{x}}_{h-1}(s,1) = \underline{\tilde{x}}_{h-1}(s) = \tilde{X}_{h-1} \underline{e}_{h-1}(s)$ may be expressed in terms of the vectors of $B(s,1)$ as

$$\underline{\tilde{x}}_{h-1}(s) = \sum_{j:\varepsilon_j \leq h-1} c_{\varepsilon_j}(s) \underline{x}_{\varepsilon_j}(s) \tag{5.38a}$$

where $c_{\varepsilon_j}(s) \in \mathbb{R}[s]$ and $\deg\{c_{\varepsilon_j}(s)\} \leq h-1-\varepsilon_j$. The above relationship then implies that

$$\mathrm{sp}\{\tilde{X}_{h-1}\} \subseteq \mathcal{R}_{\varepsilon_1} \oplus \ldots \oplus \mathcal{R}_{\varepsilon_\nu} \qquad (5.38b)$$

Given that $\mathrm{sp}\{\tilde{X}_{h-1}\}$ is a proper subspace of $\mathcal{R}_{\varepsilon_i}$, the above condition violates the completeness property of the basis $\mathcal{B}(s,\hat{s})$ and thus leads to a contradition. Therefore, the vectors $\{\underline{z}^i_j, j\in\varepsilon_i\}$ are independent and given that they span $\mathcal{T}_{\varepsilon_i}$ $(G\underline{x}^i_o=0)$, they form a basis for $\mathcal{T}_{\varepsilon_i}$.

(ii) Let $\mathcal{I}_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}$. The subspaces $\mathcal{T}_{\varepsilon_i}$, $i=g+1,\ldots,p$ are clearly nontrivial $(\neq\{0\})$ and for every such subspace $G[\underline{x}^i_1,\ldots,\underline{x}^i_{\varepsilon_i}]=G\hat{X}_{\varepsilon_i}$ is a basis matrix. Assume that the set of subspaces $\{\mathcal{T}_{\varepsilon_i} : \varepsilon_i>0\}$ is linearly dependent. Then, there exist vectors $\underline{c}_{\varepsilon_i}\in\mathbb{R}^{\varepsilon_i}$, $i=g+1,\ldots,p$, not all of them zero such that

$$G\hat{X}_{\varepsilon_{g+1}}\underline{c}_{\varepsilon_{g+1}}+\ldots+G\hat{X}_{\varepsilon_p}\underline{c}_{\varepsilon_p}=\underline{0} \qquad (5.39a)$$

If we define by $\underline{c}=[\ldots,\underline{c}^t_{\varepsilon_i},\ldots]^t$ and $\hat{X}=[\ldots;X_{\varepsilon_i};\ldots]$, then (5.39a) yields $G\hat{X}\underline{c}=\underline{0}$. Given that $\hat{X}$ has linearly independent columns, the latter condition implies that $\underline{z}=\hat{X}\underline{c}\neq\underline{0}$ and $G\underline{z}=0$, or that

$$\underline{z}\in\mathrm{sp}\{\hat{X}\}=\hat{\mathcal{R}} \text{ and } \underline{z}\in\mathcal{N}_r\{G\} \qquad (5.39b)$$

Given that $\hat{\mathcal{R}}\subset\mathcal{R}^*$, conditions (5.39b) imply that $\underline{z}\in\mathcal{N}_r\{G\}\cap\mathcal{R}^*$ and by Corollary (5.6), we have that $\underline{z}\in\mathcal{R}_\ell$. The conditions $\underline{z}\in\hat{\mathcal{R}}$ and $\underline{z}\in\mathcal{R}_\ell$ imply that $\hat{\mathcal{R}}\cap\mathcal{R}_\ell\neq\{0\}$, since $\underline{z}\neq 0$. However, by construction, the subspaces $\mathcal{R}_\ell$ and $\hat{\mathcal{R}}$ are linearly independent and thus $\hat{\mathcal{R}}\cap\mathcal{R}_\ell=\{0\}$ from which $\underline{z}=0$; given that $\hat{X}$ has full column rank, it follows that $\underline{c}=\underline{0}$ which contradicts the linear dependence assumption.

(iii) The subspace $\mathcal{T}^*$ is well defined since the set of subspaces $\{\mathcal{T}_{\varepsilon_i} : \varepsilon_i>0\}$ is linearly independent. Given that $\mathcal{T}^*$ is the image in $\mathbb{R}^m$ of the invariant subspace $\mathcal{R}^*$, it is also invariant. □

The subspace $\mathcal{T}^*$ is the G-image of $\mathcal{R}^*$ and shall be called the maximal G-right annihilating space of $sF-\hat{s}G$. In a similar way we may define the subspace $\mathcal{W}^*$ as the F-image of $\mathcal{R}^*$; $\mathcal{W}^*$ will be called the maximal F-right annihilating space of sF-ŝG and its properties are similar to those of $\mathcal{T}^*$ described by Corollary (5.6). The bases for $\mathcal{R}^*$ and $\mathcal{T}^*$ defined by the vector coefficients of a minimal basis $\mathcal{B}(s,\hat{s})$ of $\mathcal{N}_r\{sF-\hat{s}G\}$ by

$$\mathcal{B}_{\mathcal{R}*}=\{\ldots;\underline{x}_o^i,\underline{x}_1^i,\ldots,\underline{x}_{\varepsilon_i}^i;\ldots:\ i\in\underline{p}\} \tag{5.40a}$$

$$\mathcal{B}_{\mathcal{T}*}=\{\ldots;G\underline{x}_1^j,G\underline{x}_2^j,\ldots,G\underline{x}_{\varepsilon_j}^j;\ldots:\ j=g+1,\ldots,p\} \tag{5.40b}$$

will be referred to as canonical bases of $\mathcal{R}^*$ and $\mathcal{T}^*$ respectively induced by the minimal basis $\mathcal{B}(s,\hat{s})$; the corresponding basis matrices defined by $B_{\mathcal{R}*}=[\ldots;\underline{x}_o^i,\underline{x}_1^i,\ldots,\underline{x}_{\varepsilon_i}^i;\ldots]$, $i\in\underline{p}$, $B_{\mathcal{R}*}=[\ldots;G\underline{x}_1^j,G\underline{x}_2^j,\ldots,G\underline{x}_{\varepsilon_j}^j;\ldots]$, $j=g+1,\ldots,p$ will be referred to with the same name. The importance of such bases in the canonical reduction of a singular pencil is demonstrated by the following result, which is an extension of a standard result in [Gant. -1].

Proposition (5.11): Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil, $\mathcal{I}_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}$ be the set of c.m.i., $\mathcal{B}(s,\hat{s})$ be a minimal basis for $\mathcal{N}_r\{sF-\hat{s}G\}$ and let $B_{\mathcal{R}*},B_{\mathcal{T}*}$ be the canonical basis matrices of $\mathcal{R}^*,\mathcal{T}^*$ respectively induced by $\mathcal{B}(s,\hat{s})$. Then,

(i) For every pair (R,Q), where $R=[B_{\mathcal{T}*},B'']\in\mathbb{R}^{m\times m}$, $Q=[B_{\mathcal{R}*},B']\in\mathbb{R}^{n\times n}$, $|R|,|Q|\neq 0$, but with B'',B' otherwise arbitrary, then

$$R^{-1}(sF-\hat{s}G)Q=\left[\begin{array}{c|c} L_\varepsilon(s,\hat{s}) & sC-\hat{s}D \\ \hline 0 & sA-\hat{s}E \end{array}\right] \tag{5.41}$$

where $L_\varepsilon(s,\hat{s})$ has the canonical structure (5.15), and $\mathcal{N}_r\{sA-\hat{s}E\}=\{0\}$.

(ii) There always exists a choice of the B',B'' such that for the resulting (R,Q), we have

$$R^{-1}(sF-\hat{s}G)Q=\left[\begin{array}{c|c} L_\varepsilon(s,\hat{s}) & 0 \\ \hline 0 & sA-\hat{s}E \end{array}\right] \tag{5.42}$$

Proof

For the matrices $R=[B_{\mathcal{T}*},B'']$, $Q=[B_{\mathcal{R}*},B']$ we have that

$$FQ = [FB_{\mathbb{R}*},FB'] = [B_{\mathcal{T}*},B''] \begin{bmatrix} L_\varepsilon & C \\ 0 & A \end{bmatrix} \triangleq F \tag{5.43a}$$

$$GQ = [GB_{\mathbb{R}*},GB'] = [B_{\mathcal{T}*},B''] \begin{bmatrix} \hat{L}_\varepsilon & D \\ 0 & E \end{bmatrix} \triangleq G \tag{5.43b}$$

where the matrices $L_\varepsilon,\hat{L}_\varepsilon$ are defined by $L_\varepsilon(s,\hat{s})=sL_\varepsilon-\hat{s}L_\varepsilon$, where $L_\varepsilon(s,\hat{s})$ has the canonical structure (5.15). Conditions (5.43) readily follow from the properties of $B_{\mathcal{R}*}$. Clearly then $R^{-1}(sF-\hat{s}G)Q$ has the (5.41) structure. To prove that $\mathcal{N}_r\{sA-\hat{s}E\}=\{0\}$, set $\hat{s}=1$ in (5.41a) and note that the nonzero columns of $L_\varepsilon(s,1)=L_\varepsilon(s)$ define a minimal basis for the maximal $\mathbb{R}[s]$-module of $\mathbb{R}^{\Sigma\varepsilon_i}(s)$; thus, we may write

$$sC-D = L_\varepsilon(s)K(s) \tag{5.44a}$$

where $K(s)$ is some appropriate matrix with elements from $\mathbb{R}[s]$. If now $\underline{x}(s)$ is a nonzero polynomial vector for which $(sA-\hat{s}E)\underline{x}(s)=0$, then a polynomial vector $\underline{y}(s)$ may be defined by

$$\underline{y}(s) = -K(s)\underline{x}(s)+\underline{p}(s),\ \underline{p}(s)\in\mathcal{N}_r\{L_\varepsilon(s)\} \tag{5.44b}$$

By conditions (5.44a) and (5.44b), it is readily seen that

$$(s\tilde{F}-\tilde{G})\underline{r}(s) = \begin{bmatrix} L_\varepsilon(s) & sC-D \\ 0 & sA-E \end{bmatrix} \begin{bmatrix} \underline{y}(s) \\ \underline{x}(s) \end{bmatrix} = \underline{0} \tag{5.44c}$$

The polynomial vector $\underline{r}'(s)=Q^{-1}\underline{r}(s)\in\mathcal{N}_r\{sF-G\}$ and because $\underline{x}(s)\neq\underline{0}$, the corresponding supporting subspace is independent from $\mathcal{R}^*$; clearly, this contradicts the maximality property of $\mathcal{R}^*$.

(ii) The proof of part (ii) follows along similar lines to those discussed in [Gant. -1] (Vol.2, pp33-34). □

The above result demonstrates the nature of the matrix Q in the parametrisation of the basis matrices of right annihilating polynomials of $sF-\hat{s}G$ (Proposition (5.9)). Using Proposition (5.11), Proposition (5.9) may be stated as follows:

<u>Proposition (5.12)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,s]$ be a right singular pencil, $\mathcal{I}_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}$ be the set of c.m.i., $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}),\ i\in\underset{\sim}{p}\}$ be an ordered minimal basis for $\mathcal{N}_r\{sF-\hat{s}G\}$ and let

$$B_{\mathbb{R}^*} = [B_{\varepsilon_1};\ldots;B_{\varepsilon_i};\ldots;B_{\varepsilon_p}],\ B_{\varepsilon_i} = [\underline{x}_o^i,\underline{x}_1^i,\ldots,\underline{x}_{\varepsilon_p}^i] \tag{5.45}$$

be the canonical basis matrix of $\mathcal{R}^*$ induced by $\mathcal{B}(s,\hat{s})$. The basis matrix of a general k-th right annihilating vector $\underline{x}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s})$ of $sF-\hat{s}G$ may be expressed by

$$X_{k-1} = B_{\mathcal{R}^*}D_k = B_{\mathcal{R}^*}\begin{bmatrix} D_{k,o,g} \\ D_{k,\varepsilon_g+1} \\ \vdots \\ D_{k,\varepsilon_p} \end{bmatrix} = \sum_{i=1}^{g} B_{\varepsilon_i}\underline{d}^t_{k,i,o} + \sum_{i=g+1}^{p} B_{\varepsilon_j}D_{k,\varepsilon_i} \tag{5.46}$$

where $\underline{d}^t_{k,i,o}$ are k-th dimension row vectors, $D_{k,\varepsilon_i}=0_{\varepsilon_i+1,k}$ if $k<\varepsilon_i+1$ and $D_{k,\varepsilon_i}\in\mathbb{R}^{(\varepsilon_i+1)\ k}$ are the canonical blocks of (5.18) if $k\geq\varepsilon_i+1$. □

The above description of the basis matrix may be translated in terms of the corresponding homogeneous polynomials as follows:

<u>Remark (5.9)</u>: Let $\mathcal{B}(s,\hat{s})=\{\underline{x}_{\varepsilon_i}(s,\hat{s}),\ i\in\underset{\sim}{p},\ \deg\underline{x}_{\varepsilon_i}(s,\hat{s})=\varepsilon_i\}$ be a homogeneous minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$ and let $\underline{x}(s,\hat{s})\in\mathcal{N}_r\{sF-\hat{s}G\}$ be a homogeneous polynomial vector of degree k-1. Then $\underline{x}(s,\hat{s})$ may be uniquely expressed as

$$\underline{x}(s,\hat{s}) = \sum_{i:\varepsilon_i\leq k-1}\alpha_i(s,\hat{s})\underline{x}_{\varepsilon_i}(s,\hat{s}) \tag{5.47}$$

where $\alpha_i(s,\hat{s})\in\mathbb{R}[s,\hat{s}]$ are homogeneous polynomials such that $\deg\{\alpha_i(s,\hat{s})\}\leq k-1-\varepsilon_i$. □

Every k-th right annihilating polynomial vector $\underline{x}_k(s,\hat{s})$ is generated by a k-th right annihilating vector $\underline{z}_k \in \mathcal{N}_r\{T_k(F,G)\}=\mathcal{N}_r^k$; thus, the study of the $\underline{x}_k(s,\hat{s})$ vectors, as well as of the associated right annihilating spaces is intimately related to the study of basis matrices of $\mathcal{N}_r^k$. Thus, let us assume that $\mathcal{N}_r^k \neq \{0\}$ and that $N_k \in \mathbb{R}^{kn\times\theta_k}$ be a basis matrix for $\mathcal{N}_r^k$. Let us partition N_k as

$$N_k = [\underline{z}_k^1, \underline{z}_k^2, \dots, \underline{z}_k^{\theta_k}] = \begin{bmatrix} \underline{x}_{k-1}^1 & \underline{x}_{k-1}^2 & \dots & \underline{x}_{k-1}^{\theta_k} \\ \underline{x}_{k-2}^1 & \underline{x}_{k-2}^2 & \dots & \underline{x}_{k-1}^{\theta_k} \\ \vdots & \vdots & & \vdots \\ \underline{x}_1^1 & \underline{x}_1^2 & \dots & \underline{x}_1^{\theta_k} \\ \underline{x}_o^1 & \underline{x}_o^2 & \dots & \underline{x}_o^{\theta_k} \end{bmatrix} \tag{5.48}$$

and define the set of k-th order right annihilating polynomials $\mathcal{P}_k[(s,\hat{s});N_k]=\{\underline{x}_k^i(s,\hat{s}): \underline{x}_k^i(s,\hat{s})=\underline{x}_o^i\hat{s}^{k-1}+\dots+\underline{x}_{k-2}^i\hat{s}s^{k-2}+\underline{x}_{k-1}^i s^{k-1} \in \mathbb{R}^n[s,\hat{s}]$, $i\in\underset{\sim}{\theta}_k$. The set $\mathcal{P}_k[(s,\hat{s});N_k]$ will be referred to as <u>N_k-right annihilating set</u> of $sF-\hat{s}G$. The $\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$-modules generated by the polynomial vectors of $\mathcal{P}_k[(s,1);N_k]$ (set $\hat{s}=1$), $\mathcal{P}_k[(1,\hat{s});N_k]$ (set $s=1$) will be denoted by $\mathcal{M}[N_k]$, $\hat{\mathcal{M}}[N_k]$ respectively and shall be referred to as <u>$\mathbb{R}[s]$-</u>, <u>$\mathbb{R}[\hat{s}]$-</u> <u>N_k generated modules</u> of (F,G) respectively. In the following the case of $\mathcal{M}[N_k]$ modules will be considered, and the case of $\hat{\mathcal{M}}[N_k]$ modules is similar.

<u>Theorem (5.4)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil and let σ_1 be the smallest integer for which $\mathcal{N}_r^{\sigma_1} \neq \{0\}$, $\dim\mathcal{N}_r^{\sigma_1}=\rho_1$. Let N_{σ_1} be a basis matrix for $\mathcal{N}_r^{\sigma_1}$, $\mathcal{P}_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$ be the associated N_{σ_1}-right annihilating set and $\mathcal{M}[N_{\sigma_1}]$, $\hat{\mathcal{M}}[N_{\sigma_1}]$ be the corresponding modules. Then,

(i) The set $\mathcal{P}_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$ is a ρ_1-rank complete set.

(ii) $\mathcal{P}_{\sigma_1}[(s,1);N_{\sigma_1}]$, $\mathcal{P}_{\sigma_1}[(1,\hat{s});N_{\sigma_1}]$ are minimal bases for the ρ_1-rank maximal modules $\mathcal{M}[N_{\sigma_1}]$, $\hat{\mathcal{M}}[N_{\sigma_1}]$ respectively.

(iii) The modules $\mathcal{M}[N_{\sigma_1}]$, $\hat{\mathcal{M}}[N_{\sigma_1}]$ are invariants of $\mathcal{N}_r^{\sigma_1}$ and $sF-\hat{s}G$ has

$\varepsilon_1=\sigma_1-1$ as the smallest minimal index with multiplicity ρ_1.

Proof

To prove the result, it is sufficient, as well as necessary, to prove it for the $\mathcal{P}_{\sigma_1}[(s,1);N_{\sigma_1}]$ set. Thus,

(i) Assume that $\underline{x}^h_{\sigma_1}(s)$ be the first vector that is dependent on the preceding ones in $\mathcal{P}_{\sigma_1}[(s,1);N_{\sigma_1}]=\{\underline{x}^1_{\sigma_1}(s),\dots,\underline{x}^{h-1}_{\sigma_1}(s),\underline{x}^h_{\sigma_1}(s),\dots\}$. Given that σ_1 is the smallest integer for which $\mathcal{N}_r^{\sigma_1}\neq\{0\}$, it follows that $\underline{x}^j_{\sigma_1-1}\neq 0$ for $\forall\ j\in\rho_1$ (otherwise there would exist a smaller than σ_1 integer, say σ_1' for which $\mathcal{N}_r^{\sigma_1'}\neq\{0\}$). Thus all vectors $\underline{x}^j_{\sigma}(s)$ have degree $\sigma_1-1=\varepsilon_1$ and thus we may write

$$\underline{x}^h_{\sigma_1}(s)=\sum_{i=1}^{h-1}\alpha_i\underline{x}^i_{\sigma_1}(s) \tag{5.49a}$$

where $\alpha_i\in\mathbb{R}$ and not all of them zero. By equating coefficients of equal powers in (5.49 a), we have that

$$N_{\sigma_1}\underline{\alpha}=0,\ \underline{\alpha}=[-\alpha_1,\dots,-\alpha_{h-1},1,0\dots0]^t\quad. \tag{5.49b}$$

However, N_{σ_1} has linearly independent columns and thus $\underline{\alpha}=\underline{0}$, which contradicts the linear dependence assumption. Therefore $\mathcal{P}_{\sigma_1}[(s,1);N_{\sigma_1}]$ and hence $\mathcal{P}_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$, is linearly independent, and thus it is a ρ_1-rank set. By Lemma (5.3) and the minimality of σ_1, it is readily shown that all vectors in $\mathcal{P}_{\sigma_1}[(s,s);N_{\sigma_1}]$ are prime. To prove the completeness we have to show that the set of vectors

$$S_{N_{\sigma_1}}=\{\dots;\underline{x}^i_o,\underline{x}^i_1,\dots,\underline{x}^i_{\sigma_1-2},\underline{x}^i_{\sigma_1-1};\dots\ i\in\underset{\sim}{\rho_1}\}$$

is linearly independent. Assume that they are linearly dependent; then there exist $\alpha_{i,j}\in\mathbb{R}$, not all of them zero such that

$$\sum_{i=1}^{\rho_1}\sum_{j=0}^{\sigma_1-1}\alpha_{i,j}\underline{x}^i_j=0 \tag{5.50a}$$

Given that for $\forall\ i\in\underset{\sim}{\rho_1}$, the set $\{\underline{x}^i_o,\underline{x}^i_1,\dots,\underline{x}^i_{\sigma_1-1}\}$ satisfies conditions

(5.4a), then by multiplying (5.50a) by F and using (5.4a) we have that a vector $\underline{z}_1$ may be defined by

$$\underline{z}_1 = \sum_{i=1}^{\rho_1} \sum_{j=1}^{\sigma_1 -1} \alpha_{i,j-1}\underline{x}_j^i \text{ and } G\underline{z}_1 = 0 \qquad (5.50b)$$

By deploying steps similar to those used in the proof of part (ii) of Corollary (5.6), it follows that vectors of the type

$$\underline{z}_k = \sum_{i=1}^{\rho_1} \sum_{j=k}^{\sigma_1 -1} \alpha_{i,j-k}\underline{x}_j^i,\ k=1,\dots,\sigma_1 -1 \qquad (5.50c)$$

may be defined which satisfy conditions (5.4a) and thus establish the existence of a $\underline{z}(s)=\underline{z}_1+s\underline{z}_2+\dots+s^{\sigma_1 -2}\underline{z}_{\sigma_1 -1}\in N_r\{sF-\hat{s}G\}$; clearly, this violates the minimality assumption for σ_1. Therefore $S_{N_{\sigma_1}}$ is independent and $P_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$ is a ρ_1-rank complete set.

(ii) The completeness property of $P_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$ implies that the matrix $M(s)=[\underline{x}_{\sigma_1}^1(s),\dots,\underline{x}_{\sigma_1}^{\rho_1}(s)]$ has independent columns, no finite zeros and it is column reduced (otherwise the set $S_{N_{\sigma_1}}$ could be dependent. Therefore $P_{\sigma_1}[(s,1);N_{\sigma_1}]$ is a minimal basis for $M[N_{\sigma_1}]$. Since $M[N_{\sigma_1}]$ is generated by a minimal basis, it is clearly maximal.

(iii) To prove that $M[N_{\sigma_1}]$ is an invariant of $N_r^{\sigma_1}$ we have to show that it is independent of the particular basis which has been used to define it. Thus, let N'_{σ_1} be another basis of $N_r^{\sigma_1}$ and let $P'_{\sigma_1}[(s,s);N'_{\sigma_1}]$ be the corresponding set. Then $N'_{\sigma_1}=N_{\sigma_1}T, T\in\mathbb{R}^{\rho_1\times\rho_1}$, $|T|\neq 0$ and by partitioning $N'_{\sigma_1},N_{\sigma_1}$ as in (5.48), it may be readily shown that

$$\underline{x}'^{i}_{\sigma_1}(s) = \sum_{j=1}^{\rho_1} t_{ji}\underline{x}^j_{\sigma_1}(s) \qquad (5.51a)$$

where $T=[t_{ji}]$. (5.51a) then yields

$$[\underline{x}'^{1}_{\sigma_1}(s),\dots,\underline{x}'^{i}_{\sigma_1}(s),\dots,\underline{x}'^{\rho_1}_{\sigma_1}(s) =[\underline{x}^1_{\sigma_1}(s),\dots,\underline{x}^i_{\sigma_1}(s),\dots,\underline{x}^{\rho_1}_{\sigma_1}(s)]T \qquad (5.51b)$$

and thus $M[N_{\sigma_1}]=M[N'_{\sigma_1}]$. Thus $M[N_{\sigma_1}]$ is invariant. The rest of the proof is obvious.

□

The modules $M[N_{\sigma_1}],\hat{M}[N_{\sigma_1}]$ characterise the $N_r^{\sigma_1}$ space of (F,G), shall be denoted simply by $M_{\sigma_1},\hat{M}_{\sigma_1}$ and shall be referred to as the $\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$-σ_1-right annihilating modules of (F,G) respectively.

Remark (5.10): Let $N^i_{\sigma_1}$, i=1,2 be two basis matrices of $N_r^{\sigma_1}$ and let $P^i_{\sigma_1}[(s,\hat{s});N^i_{\sigma_1}]$ be the basis matrices of the associated sets $P^i_{\sigma_1}[(s,\hat{s});N^i_{\sigma_1}]$. If $N^2_{\sigma_1}=N^1_{\sigma_1}T$, $T\in\mathbb{R}^{\rho_1\times\rho_1}$, $|T|\neq 0$, then $P^2_{\sigma_1}[(s,\hat{s});N^2_{\sigma_1}]=P^1_{\sigma_1}[(s,\hat{s});N^1_{\sigma_1}]T$ and vice versa. □

A matrix $N_k\in\mathbb{R}^{kn\times\theta_k}$, which has been partitioned according to (5.48) will be referred to as naturally partitioned. With the naturally partitioned matrix N_k we may associate the vector set $S_{N_k}=\{\ldots;\underline{x}^i_{k-1},\ldots,\underline{x}^i_1,\underline{x}^i_o,\ldots:i\in\underset{\sim}{\theta_k}\}$; S_{N_k} will be referred to as the $\mathbb{R}^n$-basis set of N_k and the subset of S_{N_k}, $S^i_{N_k}=\{\underline{x}^i_{k-1},\ldots,\underline{x}^i_1,\underline{x}^i_o\}$ will be called the i-th $\mathbb{R}^n$-basis subset of N_k. N_k will be called prime if for $\forall\ i\in\underset{\sim}{\theta_k}$, $S^i_{N_k}$ is linearly independent and shall be called complete if S_{N_k} is linearly independent. Within this terminology, Theorem (5.4) yields:

Remark (5.11): If σ_1 is the minimal index for which $N_r^{\sigma_1}\neq\{0\}$ and N_{σ_1} is a basis matrix for $N_r^{\sigma_1}$, then N_{σ_1} may be naturally partitioned and it is a complete matrix.

Note that completeness of N_k implies primeness, but not vice versa. It is worth pointing out that the arguments used in the proof of Theorem (5.4) are independent from the use of the Kronecker form and thus may provide the basis for the computation of the I_c(F,G) set as well as the minimal bases of $N_r\{sF-\hat{s}G\}$ in geometric terms. Before proceeding with this study we define the following:

Let $\underline{z}=[\underline{x}^t_{k-1},\ldots,\underline{x}^t_1,\underline{x}^t_o]^t\in\mathbb{R}^{kn}$ be a naturally partitioned vector. The matrix defined by

$$T^i_{k,n}(\underline{z}) \triangleq \begin{bmatrix} \underline{x}_{k-1} & 0 & \cdots & & 0 \\ \underline{x}_{k-2} & \underline{x}_{k-1} & 0 & & \vdots \\ \underline{x}_{k-3} & \underline{x}_{k-2} & \underline{x}_{k-1} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & & \underline{x}_{k-1} \\ \underline{x}_1 & \underline{x}_2 & \underline{x}_3 & \ddots & \vdots \\ \underline{x}_o & \underline{x}_1 & \underline{x}_2 & \ddots & \vdots \\ 0 & \underline{x}_o & \underline{x}_1 & \ddots & \vdots \\ \vdots & 0 & \underline{x}_o & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \underline{x}_2 \\ \vdots & \vdots & & \ddots & \underline{x}_1 \\ 0 & 0 & \cdots & 0 & \underline{x}_o \end{bmatrix} \in \mathbb{R}^{(kn+i-1)\times i} \qquad i=1,2,\ldots \tag{5.52a}$$

will be called the <u>i-th Toeplitz matrix</u> of $\underline{z}$. If $N_k=[\underline{z}_1,\ldots,\underline{z}_j,\ldots,z_\mu]\in \mathbb{R}^{kn\times\mu}$ is a naturally partitioned matrix, then we define as the <u>i-th Toeplitz matrix</u> of N_k the matrix

$$T^i_{k,n}(N_k) \triangleq [T^i_{k,n}(\underline{z}_1);\ldots;T^i_{k,n}(\underline{z}_j);\ldots;T^i_{k,n}(\underline{z}_\mu)] \in \mathbb{R}^{(kn+i-1)\times\mu i} \tag{5.52b}$$

It is clear that for i=1, $T^1_{k,n}(N_k)=N_k$ and that for $\forall\, i$, $T^i_{k,n}(N_k)$ is naturally partitioned. Note that the definitions given for basis matrix N_k of $\mathcal{N}^k_r$, such as the sets $\mathcal{P}_k[(s,\hat{s});N_k]$ and the modules $\mathcal{M}[N_k],\hat{\mathcal{M}}[N_k]$ generated by $\mathcal{P}_k[(s,1);N_k],\mathcal{P}_k[(1,\hat{s});N_k]$ vector sets respectively, stem from the natural partitioning of N_k and thus hold true for every naturally partitioned matrix. The only distinguishing feature for the case of basis matrices of $\mathcal{N}^k_r$ is that the vectors of $\mathcal{P}_k[(s,\hat{s});N_k]$ and the modules $\mathcal{M}[N_k],\hat{\mathcal{M}}[N_k]$ possess the right annihilation property for $sF-\hat{s}G$. Of special interest here are the basis matrices N_k of $\mathcal{N}^k_r$. Some further definitions and properties of naturally partitioned matrices are discussed first.

With an $N_k\in\mathbb{R}^{kn\times\mu}$ naturally partitioned matrix, we may define the following subspaces of $\mathbb{R}^n$: $\mathcal{W}^i(N_k)\overset{\Delta}{=}\mathrm{sp}\{S^i_{N_k}\}$, $\mathcal{W}(N_k)\overset{\Delta}{=}\mathrm{sp}\{S_{N_k}\}$, where $S^i_{N_k}, S_{N_k}$ are the i-th $\mathbb{R}^n$-basis set and $\mathbb{R}^n$-basis set of N_k correspondingly. $\mathcal{W}^i(N_k), \mathcal{W}(N_k)$ will be referred to as the i-th supporting space, supporting space of N_k respectively. Clearly, $\mathcal{W}(N_k)=\sum_{i=1}^{\mu}\mathcal{W}^i(N_k)\subseteq\mathbb{R}^n$. Some useful properties of general, naturally partitioned matrices are discussed next.

Lemma (5.5): Let $N_k\in\mathbb{R}^{kn\times\mu}$ be a naturally partitioned matrix, $\mathcal{A}=\{N'_k : N'_k=N_kT, T\in\mathbb{R}^{\mu\times\mu}, |T|\neq 0\}$, $T^i_{k,n}(N_k)$ be the i-th Toeplitz matrix of N_k. The following properties hold true:

(i) If $\mathcal{W}(N_k)$ is the supporting space of N_k, then for every $N'_k\in\mathcal{A}$ and every i, i=1,2,...

$$\mathcal{W}(T^i_{k,n}(N'_k)) = \mathcal{W}(N_k) \tag{5.53}$$

(ii) If $\mathcal{M}[N_k], \hat{\mathcal{M}}[N_k]$ are the $\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$- N_k-generated modules, then for every $N'_k\in\mathcal{A}$ and every i, i=1,2,...

$$\mathcal{M}[N_k]=\mathcal{M}[T^i_{k,n}(N'_k)] \text{ and } \hat{\mathcal{M}}[N_k]=\hat{\mathcal{M}}[T^i_{k,n}(N'_k)] \tag{5.54}$$

(iii) If N_k is complete, then

(a) Every $N'_k\in\mathcal{A}$ is complete.

(b) $\mathrm{rank}[T^i_{k,n}(N_k)]=i\mu$ for all i=1,2,...

(c) The modules $\mathcal{M}[N_k], \hat{\mathcal{M}}[N_k]$ are maximal Noetherian modules.

Proof

(i) The proof readily follows from the definition of $\mathcal{W}(N_k)$ and by inspection of the special structure of the Toeplitz matrices and the relationship $N'_k=N_kT$, $T\in\mathbb{R}^{\mu\times\mu}$, $|T|\neq 0$.

(ii) The proof of $\mathcal{M}[N_k]=\mathcal{M}[N'_k]$, $N'_k\in\mathcal{A}$ is identical to that given in part (iii) of the proof of Theorem (5.4). Note that in that proof the arguments used are independent from the completeness property of N_{σ_1}.

To complete the proof we have to show that $\mathcal{M}[N_k]=\mathcal{M}[T^i_{k,n}(N_k)]$. If $\mathcal{P}_k\left[(s,\hat{s});N_k\right]=\{\underline{x}_j(s,\hat{s}),j\in\underset{\sim}{\mu}, \deg \underline{x}_j(s,\hat{s})=k\}$ is the N_k-polynomial set, then $\mathcal{P}_k[(s,1);N_k]=\{\underline{x}_j(s),j\in\underset{\sim}{\mu}, \deg \underline{x}_j(s)=k\}$ and $\mathcal{P}_{k+i}[(s,1);T^i_{k,n}(N_k)]=$ $=\{\ldots;\{\underline{x}_j(s),s\underline{x}_j(s),\ldots,s^i\underline{x}_j(s)\};\ldots,j\in\underset{\sim}{\mu}\}$. By inspection of the above two sets, it is readily seen that $\mathcal{M}[N_k]$ and $\mathcal{M}[T^i_{k,n}(N_k)]$ have the same rank and that every vector in $\mathcal{M}[N_k]$ may be expressed in terms of elements of $\mathcal{M}[T^i_{k,n}(N_k)]$ and vice versa; thus $\mathcal{M}\left[N_k\right]=\mathcal{M}[T^i_{k,n}(N_k)]$.

(iii) From the definition of completeness and given that $N_k'=N_kT$ we have that

$$N_k' = N_kT \iff \begin{bmatrix} X'_{k-1} \\ \vdots \\ X'_1 \\ X'_o \end{bmatrix} = \begin{bmatrix} X_{k-1} \\ \vdots \\ X_1 \\ X_o \end{bmatrix} T \tag{5.55}$$

Given that $\dim \operatorname{sp}\{X_j\}=\mu$, $\operatorname{sp}\{X_j'\}=\operatorname{sp}\{X_jT\}=\operatorname{sp}\{X_j\}$, $j=0,1,..,k-1$ and that $\mathcal{W}(N_k)=\overset{k}{\underset{j=0}{\oplus}} \operatorname{sp}\{X_j\}=\overset{k}{\underset{j=0}{\oplus}} \operatorname{sp}\{X_j'\}$ with $\dim \mathcal{W}(N_k)=k\mu$, part (a) is proved. Part (b) readily follows from the structure of Toeplitz matrices and the completeness of N_k. The proof of part (c) is identical to that given in the proof of Theorem (5.4). □

Using the above Lemma we may state the following result.

Corollary (5.7): Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil, σ_1 be the smallest index for which $\mathcal{N}_r^{\sigma_1}\neq\{0\}$, $\dim \mathcal{N}_r^{\sigma_1}=\rho_1$, $N_{\sigma_1}\in\mathbb{R}^{\sigma_1 n\times\rho_1}$ be a basis matrix for $\mathcal{N}_r^{\sigma_1}$ and let $\dim \mathcal{N}_r^k=\theta_k$. The following properties hold true:

(i) For all indices k: $k=\sigma_1-1+i$, $i=1,2,\ldots$, then $\theta_k\geq\rho_1(k-\sigma_1+1)$.

(ii) If σ_2 is the minimal of the indices k: $k=\sigma_1+i-1$, $i=1,2,..$ for which $\theta_k>\rho_1(k-\sigma_1+1)$, then the Toeplitz matrices $T^i_{\sigma_1,n}(N_{\sigma_1})\in\mathbb{R}^{(\sigma_1+i-1)n\times\rho_1 i}$, $i\in\underset{\sim}{\sigma_2-\sigma_1}$ are basis matrices for $\mathcal{N}_r^k$, $k=\sigma_1,\sigma_1+1,\ldots,\sigma_2-1$.

(iii) For all indices k: $k=\sigma_1-1+i$, $i=\sigma_2-\sigma_1+j$, $j=1,2,\ldots$, the basis matrices N_k of $\mathcal{N}_r^k$ may be expressed as $N_k=[T^i_{\sigma_1,n}(N_{\sigma_1});P_j]$, where P_j is an appropriate matrix and $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ is the corresponding Toeplitz matrix.

Proof

For $k=\sigma_1$, N_{σ_1} is a basis matrix and by Remark (5.11) is complete. By Lemma (5.5), part (iii)(b), $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ has full rank for $\forall i$, $i=1,2,\ldots$; it is readily verified that for $\forall k$: $k=\sigma_1-1+i$ the columns of $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ are in $\mathcal{N}^{k}_{r}$ and thus $\theta_k \geq \rho_1(k-\sigma_1+1)=\rho_1 i=\#$ columns of $T^{i}_{\sigma_1,n}(N_{\sigma_1})$. For all indices k for which $\theta_k=\rho_1(k-\sigma_1+1)=\rho_1 i$, the matrices $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ are basis matrices for $\mathcal{N}^{k}_{r}$ and this completes the proof of parts (i) and (ii). From the above arguments it follows that for all indices k: $k=\sigma_1-1+i$, $i=\sigma_2-\sigma_1+j$, $j=1,2,\ldots$, the columns of $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ are linearly independent vectors in $\mathcal{N}^{k}_{r}$, but since $\theta_k>\rho_1(k-\sigma_1+1)=\rho_1 i$, they do not span $\mathcal{N}^{k}_{r}$ any more. Therefore, starting from $T^{i}_{\sigma_1,n}(N_{\sigma_1})$ we may always expand it to a basis matrix of $\mathcal{N}^{k}_{r}$ by finding a matrix P_j, whose columns are in $\mathcal{N}^{k}_{r}$ and they are independent from the columns of $T^{i}_{\sigma_1,n}(N_{\sigma_1})$; this completes the proof. □

The above result suggests a procedure for computing basis matrices for $\mathcal{N}^{k}_{r}$ with the minimal amount of effort; that is by computing basis matrices for those spaces $\mathcal{N}^{k}_{r}$, which correspond to the singular indices of the RPAPS, and then by expanding them using the appropriate Toeplitz matrices. This technique is discussed next, and the suggested procedure leads to a geometric definition of minimal bases. We first state the following result.

Proposition (5.13): Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,s]$ be a right singular pencil, $\mathcal{N}^{k}_{r}$ be the k-th right characteristic space, $\theta_k=\dim\mathcal{N}^{k}_{r}$, and let N_k be a basis matrix for $\mathcal{N}^{k}_{r}$. Let us further assume that (σ_i,ρ_i) be the ordered pairs of integers defined by

$$\sigma_1=\min\{k:\ \mathcal{N}^{k}_{r}\neq\{0\}\},\quad \rho_1=\dim\mathcal{N}^{\sigma_1}_{r}$$
$$\sigma_i=\min\{k:\ \theta_k>\sum_{j=1}^{i-1}\rho_j(k-\sigma_j+1)\},\quad \rho_i=\theta_{\sigma_i}-\sum_{j=1}^{i-1}\rho_j(\sigma_i-\sigma_j+1) \tag{5.56}$$

For every integer σ_i, there exists a family of naturally partitioned basis matrices N_{σ_i} of $\mathcal{N}_r^{\sigma_i}$, $N_{\sigma_i}=[\hat{N}_{\sigma_i},\tilde{N}_{\sigma_i}]\in\mathbb{R}^{n\sigma_i\times\theta_{\sigma_i}}$, where $\tilde{N}_{\sigma_i}\in\mathbb{R}^{n\sigma_i\times\rho_i}$ and complete, $\mathcal{W}(\hat{N}_{\sigma_i})$ is the maximal space spanned by all right-annihilating spaces, generated by right annihilating polynomials of maximal degree σ_i-2 and $\mathcal{W}(\hat{N}_{\sigma_i})\cap\mathcal{W}(\tilde{N}_{\sigma_i})=\{0\}$.

Proof

The result will be proved by induction. For i=1, N_{σ_1} is a naturally partitioned and complete matrix (see Remark (5.11) and Theorem (5.4)). By Corollary (5.7), it follows that $N_{\sigma_2-1}=T_{\sigma_1,n}^{\sigma_2-\sigma_1}(N_{\sigma_1})$ and that $N_{\sigma_2}=$ $=[T_{\sigma_1,n}^{\sigma_2-\sigma_1+1}(N_{\sigma_1});P_1]$. Clearly $T_{\sigma_1,n}^{\sigma_2-\sigma_1+1}(N_{\sigma_1})$ has full rank and the problem is reduced to proving that P_1 may be chosen to be a complete matrix. Consider N_{σ_1}, naturally partitioned as in (5.48) and define the matrices $B_{\sigma_1},\tilde{B}_{\sigma_1}$ by

$$B_{\sigma_1} = [\underline{x}_o^1,\underline{x}_1^1,\dots,\underline{x}_{\sigma_1-1}^1;\dots;\underline{x}_o^{\rho_1},\underline{x}_1^{\rho_1},\dots,\underline{x}_{\sigma_1-1}^{\rho_1}] \tag{5.57a}$$

$$\tilde{B}_{\sigma_1} = [G\underline{x}_1^1,\dots,G\underline{x}_{\sigma_1-1}^1;\dots;G\underline{x}_1^{\rho_1},\dots,G\underline{x}_{\sigma_1-1}^{\rho_1}] \tag{5.57b}$$

By Theorem (5.4) it follows that the vectors of $\mathcal{P}_{\sigma_1}[(s,\hat{s});N_{\sigma_1}]$ is the subset of the vectors of a minimal basis of $\mathcal{N}_r\{sF-\hat{s}G\}$ which corresponds to all c.m.i. with value σ_1-1; thus, by Corollary (5.6) $\tilde{B}_{\sigma_1}$ has full rank. We may now define the matrices $Q=[B_{\sigma_1},B']\in\mathbb{R}^{n\times n}$, $R=[\tilde{B}_{\sigma_1},B'']\in\mathbb{R}^{m\times m}$, $|Q|,|Q|\neq 0$, where B',B" are suitable matrices. Using similar arguments to those of Proposition (5.11), it may be readily established that $R^{-1}(sF-\hat{s}G)Q=s\tilde{F}-s\tilde{G}$, where $s\tilde{F}-\hat{s}\tilde{G}$ has the form

$$s\tilde{F}-\hat{s}\tilde{G} = \left[\begin{array}{c:c} L_{\varepsilon_1,\rho_1}(s,\hat{s}) & sC'-\hat{s}D' \\ \hdashline 0 & sF'-\hat{s}G' \end{array}\right] \tag{5.58a}$$

$$\longleftarrow \rho_1 \text{ blocks} \longrightarrow$$

where $L_{\varepsilon_1,\rho_1}(s,\hat{s}) = \text{block-diag}\{L_{\varepsilon_1}(s,\hat{s}),\dots,L_{\varepsilon_1}(s,\hat{s})\}$, $\varepsilon_1=\sigma_1-1$ and

$N_r\{sF';sG'\}$ has no vector in it with homogeneous degree less than σ_1 (this result is a trivial extension of the result stated in [Gant. - 1] (Vol.2, pp33-34)); furthermore, the matrices B',B" may be suitably chosen such that

$$s\widetilde{F}-\hat{s}\widetilde{G} = \left[\begin{array}{c|c} L_{\varepsilon_1,\rho_1}(s,\hat{s}) & 0 \\ \hline 0 & sF'-\hat{s}G' \end{array}\right] \tag{5.58b}$$

Using the special (R,Q) pair for which $s\widetilde{F}-\hat{s}\widetilde{G}$ is given by (5.58b), it follows that $N_r^{\sigma_2}(\widetilde{F},\widetilde{G})$ may be expressed as a direct sum of the null spaces $N_r^{\sigma_2}(L_{\varepsilon_1,\rho_1},\hat{L}_{\varepsilon_1,\rho_1})$ and $N_r^{\sigma_2}(F',G')$, where $L_{\varepsilon_1,\rho_1}(s,\hat{s})=sL_{\varepsilon_1,\rho_1}-\hat{s}\hat{L}_{\varepsilon_1,\rho_1}$. Let U_{σ_2} be the representation of the basis matrix $T_{\sigma_1,n}^{\sigma 2-\sigma 1+1}(N_{\sigma_1})$, when it is restricted to the $L_{\varepsilon_1,\rho_1}(s,\hat{s})$ subpencil and W_{σ_2} be any basis matrix of $N_r^{\sigma_2}(F',G')$. We may partition U_{σ_2},W_{σ_2} as

$$U_{\sigma_2} = \left[\begin{array}{c} U'_{\sigma_2-1} \\ \hline U'_{\sigma_2-2} \\ \hline \vdots \\ \hline U'_o \end{array}\right] \begin{array}{c} \updownarrow \rho_1\sigma_1 \\ \updownarrow \rho_1\sigma_1 \\ \\ \updownarrow \rho_1\sigma_1 \end{array}, \quad W_{\sigma_2} = \left[\begin{array}{c} W'_{\sigma_2-1} \\ \hline W'_{\sigma_2-2} \\ \hline \vdots \\ \hline W'_o \end{array}\right] \begin{array}{c} \updownarrow n-\rho_1\sigma_1 \\ \updownarrow n-\rho_1\sigma_1 \\ \\ \updownarrow n-\rho_1\sigma_1 \end{array} \tag{5.59a}$$

Then, a basis matrix for $N_r^{\sigma_2}(\widetilde{F},\widetilde{G})$ is defined by

$$Z_{\sigma_2} = \left[\begin{array}{c|c} U'_{\sigma_2-1} & 0 \\ \hline 0 & W'_{\sigma_2-1} \\ \hline \vdots & \vdots \\ \hline U'_o & 0 \\ \hline 0 & W'_o \end{array}\right] = [\widetilde{U}_{\sigma_2},\widetilde{W}_{\sigma_2}] \tag{5.59b}$$

By Theorem (5.4) and Remark (5.11), W_{σ_2} is naturally partitioned and complete matrix and thus $\widetilde{W}_{\sigma_2}$ is also naturally partitioned and complete matrix. By translating the expression for Z_{σ_2} back to our original frame, we have that if $\langle Q\rangle=\text{diag}\{\underbrace{Q,\dots,Q}_{\sigma_2}\}$, then

$$N_{\sigma_2} = [T^{\sigma_2-\sigma_1+1}_{\sigma_1,n}(N_{\sigma_1});P_1] = \langle Q\rangle Z_{\sigma_2} = [\langle Q\rangle\tilde{U}_{\sigma_2},\langle Q\rangle\tilde{W}_{\sigma_2}] = [\hat{N}_{\sigma_2},\tilde{N}_{\sigma_2}] \tag{5.59c}$$

Clearly N_{σ_2} is a basis matrix for $\mathcal{N}_r^{\sigma_2}$ and $\tilde{N}_{\sigma_2} \in \mathbb{R}^{\sigma_2 n\times\rho_2}$ and has full rank. The matrix $\tilde{N}_{\sigma_2}$ is complete since the columns of

$$P_{\sigma_2} = \left[\begin{array}{c|c|c} \multicolumn{3}{c}{0} \\ \hline W'_{\sigma_2-1} & \cdots & W'_o \end{array}\right] \tag{5.59d}$$

are linearly independent (completeness of $\tilde{W}_{\sigma_2}$), and thus also the columns of QP_{σ_2} are linearly independent. By Lemma (5.5), $\mathcal{W}(\hat{N}_{\sigma_2})=\mathcal{W}(N_{\sigma_1})$ and it is clearly the maximal space spanned by the supporting subspaces of all right annihilating polynomials of maximal degree σ_2-2. Finally, for the matrix

$$Q\left[\begin{array}{c|c|c|c|c|c} U'_{\sigma_2-1} & \cdots & U'_o & \multicolumn{3}{c}{0} \\ \hline \multicolumn{3}{c|}{0} & W'_{\sigma_2-1} & \cdots & W'_o \end{array}\right] = Q\left[\begin{array}{c|c} P_{\sigma_2} & R_{\sigma_2} \end{array}\right] \tag{5.59e}$$

there are no nonzero vectors $\underline{\alpha},\underline{\beta}$ for which $P_{\sigma_2}\underline{\alpha}+R_{\sigma_2}\underline{\beta}=0$; therefore $\mathcal{W}(\hat{N}_{\sigma_2})\cap\mathcal{W}(\tilde{N}_{\sigma_2})=\{0\}$ and for the step i=2 the result is established. The generalization for any arbitrary index is identical. The arbitrariness of the basis matrices W_{σ_i} implies the existence of a family. □

The set of ordered pairs of integers $\mathcal{K}_r(F,G)=\{(\sigma_i,\rho_i):\sigma_1<\ldots<\sigma_i<\ldots\}$ defined by (5.56) will be called the <u>right set of singularity</u> of (F,G) and every index σ_i will be referred to as a <u>right singular index</u> of (F,G). A useful remark that readily follows from the proof of the above result (expression (5.59e) etc.) is stated next.

<u>Remark (5.12)</u>: Let $N_{\sigma_i}=[\hat{N}_{\sigma_i},\tilde{N}_{\sigma_i}]\in\mathbb{R}^{n\sigma_i\times\theta_{\sigma_i}}$ be a basis matrix for $\mathcal{N}_r^{\sigma_i}$ which satisfies the properties of Proposition (5.13). Every other basis matrix N'_{σ_i} of $\mathcal{N}_r^{\sigma_i}$ which is related to N_{σ_i} by

$$N'_{\sigma_i} = \left[\hat{N}'_{\sigma_i},\tilde{N}'_{\sigma_i}\right] = N_{\sigma_i}H = \left[\hat{N}_{\sigma_i},\tilde{N}_{\sigma_i}\right]\begin{bmatrix} H_1 & H_2 \\ 0 & H_3 \end{bmatrix}$$

where $H \in \mathbb{R}^{\theta_{\sigma_i} \times \theta_{\sigma_i}}$, $|H| \neq 0$, has also the properties of the bases defined by Proposition (5.13). □

In the following we investigate the relationship between the sets $K_r(F,G)$ and $I_c(F,G)$.

<u>Proposition (5.14)</u>: Let $K_r(F,G)=\{(\sigma_i,\rho_i), i=1,2,\dots\}$ and $I_c(F,G)=\{(\varepsilon_i,\rho_i'), i \in \underset{\sim}{\mu}\}$ be the right set of singularity and the c.m.i. sets respectively of a right singular pencil. Then $K_r(F,G)$ is finite and its cardinality is μ; furthermore, for $\forall\, i \in \underset{\sim}{\mu}$, $\rho_i=\rho_i'$ and $\sigma_i=\varepsilon_i+1$.

<u>Proof</u>

The proof of the result readily follows by Theorem (5.2) and Proposition (5.6). An alternative proof to this result will be also given at the end of this section. □

<u>Systematic Procedure for the computation of $K_r(F,G)$</u>

The right set of singularity has been defined in Proposition (5.13) and by Proposition (5.14), it is related to $I_c(F,G)$ in a simple manner and thus may be used for the computation of $I_c(F,G)$. The systematic procedure adopted for the computation of $K_r(F,G)$ is described below:

(i) Define the set of indices $I_1=\{k: N_r^k \neq \{0\}, k \in \underset{\sim}{\tilde{\sigma}}$, where $\tilde{\sigma}=m+1$, if $m<n$ and $\tilde{\sigma}=n$ if $m \geq n\}$.

  (a) If $I_1=\emptyset$, then $K_r(F,G)=\emptyset$ and $sF-\hat{s}G$ is not right singular.

  (b) If $I_1 \neq \emptyset$, find $\sigma_1=\min I_1$ and define $\rho_1=\dim N_r^{\sigma_1}=\theta_{\sigma_1}$. Define the number $\pi_{\sigma_1} \overset{\Delta}{=} n-\rho_1\sigma_1-\sigma_1$.

  (c) If $\pi_{\sigma_1} \leq 0$, then $\mu=1$ and $K_r(F,G)=\{(\sigma_1,\rho_1)\}$.

  (d) If $\pi_{\sigma_1}>0$, continue to the following step.

(ii) Define the set of indices $I_2=\{k: \theta_k>\rho_1(k-\sigma_1+1), k=\sigma_1+i, i \in \underset{\sim}{\pi}_{\sigma_1}\}$.

  (a) If $I_2=\emptyset$, then $\mu=1$ and $K_r(F,G)=\{(\sigma_1,\rho_1)\}$.

  (b) If $I_2 \neq \emptyset$, find $\sigma_2=\min I_2$ and define $\rho_2=\theta_{\sigma_2}-\rho_1(\sigma_2-\sigma_1+1)$.

Define the number $\pi_{\sigma_2} \overset{\Delta}{=} n-\rho_1\sigma_1-\rho_2\sigma_2-\sigma_2$.

(c) If $\pi_{\sigma_2} \leq 0$, then $\mu=2$ and $\mathcal{K}_r(F,G)=\{(\sigma_1,\rho_1),(\sigma_2,\rho_2)\}$.

(d) If $\pi_{\sigma_2} > 0$, continue to the following step.

(iii) Assume that from the previous steps the subset $\mathcal{K}_r^i(F,G)= =\{(\rho_1,\sigma_1),\dots,(\rho_i,\sigma_i)\}$ of $\mathcal{K}_r(F,G)$ has been defined as well as the number $\pi_{\sigma_i} \overset{\Delta}{=} n-\sum_{j=1}^{i}\rho_j\sigma_j-\sigma_i$ and that $\pi_{\sigma_i}>0$, otherwise the procedure could have terminated. Define the set of indices $\mathcal{I}_{i+1}=\{k: \theta_k > \sum_{j=1}^{i}\rho_j(k-\sigma_j+1), k=\sigma_i+p, p\in\underset{\sim}{\pi}_{\sigma_i}\}$.

(a) If $\mathcal{I}_{i+1}=\emptyset$, then $\mu=i$ and $\mathcal{K}_r^i(F,G)=\mathcal{K}_r(F,G)$.

(b) If $\mathcal{I}_{i+1}\neq\emptyset$, find $\sigma_{i+1}=\min \mathcal{I}_{i+1}$ and define the numbers $\rho_{i+1}=\theta_{\sigma_{i+1}}-\sum_{j=1}^{i}\rho_j(\sigma_{i+1}-\sigma_j+1)$ and $\pi_{\sigma_{i+1}} \overset{\Delta}{=} n-\sum_{j=1}^{i+1}\rho_j\sigma_j-\sigma_{i+1}$.

(c) If $\pi_{\sigma_{i+1}} \leq 0$, then $\mu=i+1$ and $\mathcal{K}_r(F,G)=\{(\sigma_j,\rho_j),j=1,\dots,i+1\}$.

(d) If $\pi_{\sigma_i+1}>0$, continue to the following step.

Note that $\pi_{\sigma_j}>\pi_{\sigma_j+1}$ and thus the procedure terminates either when for some τ, $\pi_{\sigma_\tau}\leq 0$, or when $\mathcal{I}_{\tau+1}=\emptyset$. Having found the set $\mathcal{K}_r(F,G)=\{(\sigma_i,\rho_i),i\in\underset{\sim}{\mu}\}$, then $\mathcal{I}_c(F,G)=\{(\sigma_i-1,\rho_i),i\in\underset{\sim}{\mu}\}$. Finally, note that σ_μ is the right index of (F,G). □

Proposition (5.13) has established the existence of basis matrices $N_{\sigma_i}=[\hat{N}_{\sigma_i},\tilde{N}_{\sigma_i}]\in\mathbb{R}^{n\sigma_i\times\theta_{\sigma_i}}$ for $\mathcal{N}_r^{\sigma_i}$ for $\forall\sigma_i$ singular index; such matrices are characterised by the properties: (1) they are naturally partitioned, (2) $\mathcal{W}(\hat{N}_{\sigma_i})$ is the maximal space spanned by all right annihilating spaces generated by vectors of maximal degree of σ_i-2, (3) $\tilde{N}_{\sigma_i}$ is complete and (4) $\mathcal{W}(\hat{N}_{\sigma_i})\cap\mathcal{W}(\tilde{N}_{\sigma_i})=\{0\}$. A basis matrix of $\mathcal{N}_r^{\sigma_i}$, where σ_i is a singular index, which satisfies the above four properties will be called *σ_i-regular basis matrix* of $\mathcal{N}_r^{\sigma_i}$ and the submatrix $\tilde{N}_{\sigma_i}\in\mathbb{R}^{\sigma_i n\times\rho_i}$ will be referred to as a *(σ_i,ρ_i)-regular basis complement of $\mathcal{N}_r^{\sigma_i}$*. The σ_i-regular basis matrices provide the means for the computation of minimal bases of $\mathcal{N}_r\{sF-\hat{s}G\}$ and they also reveal a number of important properties of $\mathcal{N}_r$ $sF-\hat{s}G$ as it is shown next. A useful family of basis matrices of the $\mathcal{N}_r^k$ subspaces is considered first.

<u>Proposition (5.15)</u>: Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil, $\mathcal{K}(F,G)=\{(\sigma_i,\rho_i),i\in\underline{\mu};\sigma_1<\ldots<\sigma_\mu\}$ be the right set of singularity and let $T^i_{\tau,n}(A_\tau)$ denote the i-th Toeplitz matrix of the naturally partitioned matrix $A_\tau\in\mathbb{R}^{\tau n\times p}$. For all $k\geq\sigma_1$, there exist families of basis matrices $\{N^*_k: k\geq\sigma_1\}$ for the sequence of subspaces $\{\mathcal{N}^k_r: k\geq\sigma_1\}$ which are defined as follows:

(i) For every k: $\sigma_1\leq k<\sigma_2$, $N^*_k=[T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1})]$, where $\tilde{N}_{\sigma_1}=N^*_{\sigma_1}=T^1_{\sigma_1,n}(\tilde{N}_{\sigma_1})$ is an arbitrary basis matrix of $\mathcal{N}_r^{\sigma_1}$.

(ii) For every k: $\sigma_i\leq k<\sigma_{i+1}$, $i\geq2$, $N^*_k=[T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1});\ldots;T^{k-\sigma_i+1}_{\sigma_i,n}(\tilde{N}_{\sigma_i})]$, where $\tilde{N}_{\sigma_i}=T^1_{\sigma_i,n}(\tilde{N}_{\sigma_i})\in\mathbb{R}^{\sigma_i n\times\rho_i}$ is a basis matrix for any arbitrary subspace $\mathcal{V}_{\sigma_i}$ for which $\mathcal{N}_r^{\sigma_i}=\mathcal{V}_{\sigma_i}$ col-sp$[T^{\sigma_i-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1});\ldots;T^{\sigma_i-\sigma_{i-1}+1}_{\sigma_i-1,n}(\tilde{N}_{\sigma_i-1})]$.

(iii) For every k: $k\geq\sigma_\mu$, $N^*_k=[T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1});\ldots;T^{k-\sigma_\mu+1}_{\sigma_\mu,n}(\tilde{N}_{\sigma_\mu})]$, where $\tilde{N}_{\sigma_\mu}=T^1_{\sigma_\mu,n}(\tilde{N}_{\sigma_\mu})\in\mathbb{R}^{\sigma_\mu n\times\rho_\mu}$ is a basis matrix for any arbitrary subspace $\mathcal{V}_{\sigma_\mu}$ for which $\mathcal{N}_r^{\sigma_\mu}=\mathcal{V}_{\sigma_\mu}\oplus$col-sp$[T^{\sigma_\mu-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1});\ldots;T^{\sigma_\mu-\sigma_{\mu-1}+1}_{\sigma_{\mu-1},n}(\tilde{N}_{\sigma_{\mu-1}})]$.

<u>Proof</u>

(i) By Remark (5.11), any basis matrix, say $\tilde{N}_{\sigma_1}$ of $\mathcal{N}^r_{\sigma_1}$ is complete and by the part (ii) of Corollary (5.7) this part of the result is established.

(ii) From part (i) it follows that $N^*_{\sigma_2-1}=T^{\sigma_2-\sigma_1}_{\sigma_1,n}(\tilde{N}_{\sigma_1})$ is a basis of $\mathcal{N}_r^{\sigma_2-1}$. The column vectors of $\hat{N}^*_{\sigma_2}=T^{\sigma_2-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1})$ are in $\mathcal{N}_r^{\sigma_2}$ and they are linearly independent; furthermore, $\mathcal{Z}^*_{\sigma_2}=$col-sp$\{\hat{N}^*_{\sigma_2}\}$ is a proper subspace of $\mathcal{N}_r^{\sigma_2}$ and thus there always exist nontrivial subspaces $\mathcal{V}^*_{\sigma_2}$ such that $\mathcal{N}_r^{\sigma_2}=\mathcal{Z}^*_{\sigma_2}\oplus\mathcal{V}^*_{\sigma_2}$. Let $\tilde{N}^*_{\sigma_2}$ be a basis matrix for $\mathcal{V}^*_{\sigma_2}$, then $N^*_{\sigma_2}=[\hat{N}^*_{\sigma_2},\tilde{N}^*_{\sigma_2}]$ is clearly a basis matrix for $\mathcal{N}_r^{\sigma_2}$. We shall prove that $N^*_{\sigma_2}$ is a σ_2-regular basis matrix. By Proposition (5.13), there exists a σ_2-regular basis matrix $N_{\sigma_2}=[\hat{N}_{\sigma_2},\tilde{N}_{\sigma_2}]$ of $\mathcal{N}_r^{\sigma_2}$. Thus, there exists an $M\in\mathbb{R}^{\theta_{\sigma_2}\times\theta_{\sigma_2}}$, $|M|\neq0$ such that

$$N^{*}_{\sigma_2} = \left[\hat{N}^{*}_{\sigma_2}, \tilde{N}^{*}_{\sigma_2}\right] = N_{\sigma_2} M = \left[\hat{N}_{\sigma_2}, \tilde{N}_{\sigma_2}\right] \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \tag{5.60a}$$

By the construction of the σ_2-regular basis matrix N_{σ_2} (eqns. (5.59a)-(5.59c)) and Lemma (5.5) it follows that col-sp$\{\hat{N}^{*}_{\sigma_2}\}$=col-sp$\{\hat{N}_{\sigma_2}\}$ and thus by (5.60a) it is readily seen that $M_3=0$ (otherwise some columns of $\tilde{N}_{\sigma_2}$ could have been linearly dependent on the columns of $\hat{N}_{\sigma_2}$). Since $M_3=0$, then by (5.60a) we have

$$N^{*}_{\sigma_2} = \left[\hat{N}^{*}_{\sigma_2}, \tilde{N}^{*}_{\sigma_2}\right] = \left[\hat{N}_{\sigma_2}, \tilde{N}_{\sigma_2}\right] \begin{bmatrix} M_1 & M_2 \\ 0 & M_4 \end{bmatrix} \tag{5.60b}$$

By Remark (5.12), $N^{*}_{\sigma_2}$ belongs to the family of σ_2-regular basis matrices of $\mathcal{N}_r^{\sigma_2}$ and thus $\tilde{N}^{*}_{\sigma_2}$ is a (σ_2,ρ_2)-regular basis complement of $\mathcal{N}_r^{\sigma_2}$. By Lemma (5.5), $T^{k-\sigma_2+1}_{\sigma_2,n}(\tilde{N}^{*}_{\sigma_2})$ has full rank for $\forall k\geq\sigma_2$ and its columns are in $\mathcal{N}_r^{k}$. From the previous step, $T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1})$ has full rank for $\forall k\geq\sigma_1$ (Lemma (5.5)), and its columns are also in $\mathcal{N}_r^{k}$. The columns of

$$N^{*}_{k} = \left[T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1});T^{k-\sigma_2+1}_{\sigma_2,n}(\tilde{N}^{*}_{\sigma_2})\right] \tag{5.60c}$$

are thus in $\mathcal{N}_r^{k}$ for $\forall k\geq\sigma_2$ and they are linearly independent; in fact, if we assume dependence of the columns of (5.60c) matrix, then it is readily shown that $\mathcal{W}(T^{k-\sigma_1+1}_{\sigma_1,n}(\tilde{N}_{\sigma_1}))\cap\mathcal{W}(T^{k-\sigma_2+1}_{\sigma_2,n}(\tilde{N}^{*}_{\sigma_2}))\neq\{0\}$ which in turn implies that $\mathcal{W}(\tilde{N}_{\sigma_1})\cap\mathcal{W}(\tilde{N}^{*}_{\sigma_2})\neq\{0\}$ and thus violates the (σ_2,ρ_2)-regular basis complement nature of $\tilde{N}_{\sigma_2}$. Thus, rank$(N^{*}_{k})=\rho_1(k-\sigma_1+1)+\rho_2(k-\sigma_2+1)=\dim\mathcal{N}_r^{k}$ for $\forall k$: $\sigma_2\leq k<\sigma_3$ and the N^{*}_{k}, defined by (5.60c), is a basis matrix for $\mathcal{N}_r^{k}$. By induction (the rest of the steps are similar) the proof of the result is readily completed.

□

Every family $\{N^{*}_{k}: k\geq\sigma_1\}$ of basis matrices for the sequence of subspaces $S_r(F,G)=\{\mathcal{N}_r^{k}: k\geq\sigma_1\}$, constructed as in Proposition (5.15) will be called a <u>complete family</u>. By the construction procedure, it is clear that a complete family is uniquely defined, as long as the set of basis matrices

$\{\tilde{N}_{\sigma_i}, i\in\underset{\sim}{\mu}\}$, for the set of subspaces $\mathcal{G}=\{\mathcal{V}_{\sigma_i}: i\in\underset{\sim}{\mu}, \mathcal{V}_{\sigma_1}=\mathcal{N}_r^{\sigma_1}\}$ has been defined. The set of subspaces $\mathcal{G}$ will be referred to as a set of generating subspaces of $\mathcal{S}_r(F,G)$ and every set of basis matrices of $\mathcal{G}$, $\mathcal{A}_{\mathcal{G}}=\{\tilde{N}_{\sigma_i}, i\in\underset{\sim}{\mu}\}$ will be called the set of generators of the complete family $\mathcal{B}(\mathcal{A}_{\mathcal{G}})=\{N_k^*: k\geq\sigma_1\}$. Note that the elements of $\mathcal{G}$ (apart from $\mathcal{V}_{\sigma_1}$) are not uniquely defined; and the family of all sets of generating subspaces of (F,G) will be denoted by $<(F,G)_r>$. For every $\mathcal{G}\in<(F,G)_r>$, there exists a family of $\mathcal{A}_{\mathcal{G}}$ sets of generators which shall be denoted by $<\mathcal{A}_{\mathcal{G}}>$. If $\{N_k: k\geq\sigma_1\}$ is a sequence of basis matrices for the elements of $\mathcal{S}_r(F,G)$ (clearly, such a sequence is not uniquely defined), then for $\forall\, N_i\in\{N_k: k\geq\sigma_1\}$ we shall denote by $\mathcal{P}[(s,\hat{s});N_i]$ the N_k-right annihilating set, by $\mathcal{M}[N_i],\hat{\mathcal{M}}[N_i]$ the corresponding $\mathbb{R}[s]$-, $\mathbb{R}[\hat{s}]$-N_i-generated modules and by $\mathcal{W}(N_i)$ the supporting subspace of N_i. The main result of this section is stated below.

Theorem (5.5): Let $sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}]$ be a right singular pencil and let $\mathcal{K}(F,G)=\{(\sigma_i,\rho_i), i\in\underset{\sim}{\mu}: \sigma_1<\ldots<\sigma_\mu\}$ be the right set of singularity, $\mathcal{B}(\mathcal{A}_{\mathcal{G}})=\{N_k^*: k\geq\sigma_1\}$ be a complete family of basis matrices of $\mathcal{S}_r(F,G)$, corresponding to a set of generators $\mathcal{A}_{\mathcal{G}}=\{\tilde{N}_{\sigma_i}, i\in\underset{\sim}{\mu}\}$ and let $\mathcal{B}=\{N_k: k\geq\sigma_1\}$ be a general family of basis matrices of $\mathcal{S}_r(F,G)$. Then,

(i) For every singular index σ_i, the matrix $N_{\sigma_i}^*\in\mathcal{B}(\mathcal{A}_{\mathcal{G}})$ is a σ_i-regular basis matrix of $\mathcal{N}_r^{\sigma_i}$ and $\tilde{N}_{\sigma_i}\in\mathcal{A}_{\mathcal{G}}$ is a (σ_i,ρ_i)-regular basis complement of $\mathcal{N}_r^{\sigma_i}$. Furthermore, for any set $\mathcal{A}_{\mathcal{G}}$, the corresponding set of supporting subspaces $\mathcal{W}(\mathcal{A}_{\mathcal{G}})=\{\mathcal{W}(\tilde{N}_{\sigma_i}), i\in\underset{\sim}{\mu}\}$ is linearly independent.

(ii) Let $\Omega_k=\{\sigma_t, t\in\underset{\sim}{\nu}: \sigma_t\in\{\sigma_1,\ldots,\sigma_\mu\}$ and $\sigma_t\leq k\}$. For any basis matrix N_k of $\mathcal{N}_r^k$, the supporting subspace $\mathcal{W}(N_k)$ is an invariant of $\mathcal{N}_r^k$ and it may be expressed as

$$\mathcal{W}(N_k) = \mathcal{W}(\tilde{N}_{\sigma_1}) \oplus \ldots \oplus \mathcal{W}(\tilde{N}_{\sigma_\nu}) \tag{5.61}$$

(iii) For every set of generators $\mathcal{A}_G$, then

(a) The set of homogeneous binary polynomials $\mathcal{P}(\mathcal{A}_G)=\{\mathcal{P}_{\sigma_i}[(s,\hat{s}):\tilde{N}_{\sigma_i}],\ i\in\underset{\sim}{\mu}\}$ is complete and of rank $\rho=\sum_{i=1}^{\mu}\rho_i$.

(b) For every singular index σ_i the $\mathbb{R}[s]$-module $\mathcal{M}[\tilde{N}_{\sigma_i}]$ has rank ρ_i, it is maximal Noetherian and the vectors of $\mathcal{P}_{\sigma_i}[(s,1):\tilde{N}_{\sigma_i}]$ define a minimal basis for $\mathcal{M}[\tilde{N}_{\sigma_i}]$. Furthermore, the set of $\mathbb{R}[s]$-modules defined by $\mathcal{M}[\mathcal{A}_G]=\{\mathcal{M}[\tilde{N}_{\sigma_i}],\ i\in\underset{\sim}{\mu}\}$ is linearly independent.

(iv) Let $\Omega_k=\{\sigma_t, t\in\underset{\sim}{\nu}:\ \sigma_t\in\{\sigma_1,\dots,\sigma_\mu\}$ and $\sigma_t\leq k\}$. For any basis matrix N_k of $\mathcal{N}_r^k$, the $\mathbb{R}[s]$-module $\mathcal{M}[N_k]$ is an invariant of $\mathcal{N}_r^k$, it is maximal Noetherian of $\sum_{t=1}^{\nu}\rho_i$ rank and it may be expressed as

$$\mathcal{M}[N_k]=\mathcal{M}[N_{\sigma_1}]\oplus\dots\oplus\mathcal{M}[N_{\sigma_\nu}] \tag{5.62}$$

<u>Proof</u>

(i) The proof that for $\forall\sigma_i$, the matrix $N^*_{\sigma_i}$ is a σ_i-regular matrix of $\mathcal{N}_r^{\sigma_i}$ is explicitly stated as part of the proof of Proposition (5.15). By construction, $N^*_{\sigma_i}$ is expressed as

$$N^*_{\sigma_i}=\left[T_{\sigma_1,n}^{\sigma_i-\sigma_1+1}(\tilde{N}_{\sigma_1});\dots;T_{\sigma_{i-1},n}^{\sigma_i-\sigma_{i-1}+1}(\tilde{N}_{\sigma_i-1});\tilde{N}_{\sigma_i}\right]=\left[\hat{N}_{\sigma_i};\tilde{N}_{\sigma_i}\right] \tag{5.63a}$$

and $\mathcal{W}(\hat{N}_{\sigma_i})\cap\mathcal{W}(\tilde{N}_{\sigma_i})=\{0\}$. By Lemma (5.5) $\mathcal{W}(\hat{N}_{\sigma_i})=\sum_{j=1}^{i-1}\mathcal{W}(\tilde{N}_{\sigma_j})$ and thus $\sum_{j=1}^{i-1}\mathcal{W}(\tilde{N}_{\sigma_j})\cap\mathcal{W}(\tilde{N}_{\sigma_i})=\{0\}$ for $\forall i\in\underset{\sim}{\mu}$; the last condition clearly implies the independence of the set $\mathcal{W}(\mathcal{A}_G)$, as well as that

$$\mathcal{W}(N^*_{\sigma_i})=\mathcal{W}(\tilde{N}_{\sigma_1})\oplus\dots\oplus\mathcal{W}(\tilde{N}_{\sigma_i}) \tag{5.63b}$$

(ii) Let N_k be any basis matrix and N^*_k the k-th element of a complete family of basis matrices. There exists a square, full rank matrix H such that $N_k=N^*_kH$. This relationship, in the natural partitioned form, may be expressed as

$$N_k = \begin{bmatrix} N_k^{k-1} \\ \hdashline N_k^{k-2} \\ \hdashline \vdots \\ \hdashline N_k^{o} \end{bmatrix} = \begin{bmatrix} N_k^{*k-1} \\ \hdashline N_k^{*k-2} \\ \hdashline \vdots \\ \hdashline N_k^{*o} \end{bmatrix} H = N_k^{*}H \qquad (5.64a)$$

and thus $\mathcal{W}(N_k)$=col-sp$\{[N_k^{k-1},\ldots,N_k^{o}]\}$=col-sp$\{[N_k^{*k-1}H,\ldots,N_k^{*o}H]\}$= =col-sp$\{[N_k^{*k-1},\ldots,N_k^{*o}]\text{diag}\{\underbrace{H,\ldots,H}_{k}\}\}=\mathcal{W}(N_k^{*})$. This condition implies that: (1) $\mathcal{W}(N_k)$ is invariant of $\mathcal{N}_r^k$ and (2) that the properties of $\mathcal{W}(N_k)$ may be studied by considering a complete basis for $\mathcal{N}_r^k$.

Let Ω_k be the set of singular indices for which $\sigma_t \leq k$. Then, N_k^{*} may be expressed as

$$N_k^{*} = \left[T_{\sigma_1,n}^{k-\sigma_1+1}(\tilde{N}_{\sigma_1});\ldots;T_{\sigma_i,n}^{k-\sigma_\nu+1}(\tilde{N}_{\sigma_\nu}) \right] \qquad (5.64b)$$

By Lemma (5.5), $\mathcal{W}(N_k^{*})=\sum_{t=1}^{\nu}\mathcal{W}(\tilde{N}_{\sigma_t})$ and by the independence of the set $\mathcal{W}(A_G)$ the result is established.

(iii) (a) Since $\tilde{N}_{\sigma_i}$ is a naturally partitioned and complete matrix then the set $\mathcal{P}_{\sigma_i}[(s,\hat{s});\tilde{N}_{\sigma_i}]$ is linearly independent and rank ρ_i, since otherwise, the completeness of $\tilde{N}_{\sigma_i}$ is violated. The primeness and the independence of the characteristic subspaces of the vectors of $\mathcal{P}_{\sigma_i}[(s,\hat{s}):\tilde{N}_{\sigma_i}]$ follow immediately from the primeness of $\tilde{N}_{\sigma_i}$. Thus, $\mathcal{P}_{\sigma_i}[(s,\hat{s});\tilde{N}_{\sigma_i}]$ is a complete set. The independence of the set of subspaces $\mathcal{W}(A_G)$ established in part (i) and the completeness of every $\mathcal{P}_{\sigma_i}[(s,\hat{s});\tilde{N}_{\sigma_i}]$ implies that the vectors of $\mathcal{P}(A_G)$ are linearly independent (otherwise the independence of $\mathcal{W}(A_G)$ is violated); clearly, the rank of $\mathcal{P}(A_G)$ is $\sum_{i=1}^{\mu}\rho_i$. The completeness property of every subset $\mathcal{P}_{\sigma_i}[(s,\hat{s});\tilde{N}_{\sigma_i}]$ of $\mathcal{P}(A_G)$ and the independence of the set $\mathcal{W}(A_G)$ implies that $\mathcal{P}(A_G)$ is complete.

(b) By definition $\mathcal{M}[\tilde{N}_{\sigma_i}]$ is generated by $\mathcal{P}_{\sigma_i}[(s,1);\tilde{N}_{\sigma_i}]$; the completeness of $\mathcal{P}_{\sigma_i}[(s,1);\tilde{N}_{\sigma_i}]$ implies that it is a minimal basis of $\mathcal{M}[\tilde{N}_{\sigma_i}]$ which has a minimal degree and thus $\mathcal{M}[\tilde{N}_{\sigma_i}]$ is maximal Noetherian and its rank is ρ_i.

The completeness of the set of vectors $\mathcal{P}(A_{\mathcal{G}})$, immediately imply the independence of the $\mathcal{M}[A_{\mathcal{G}}]$ set.

(iv) Let N_k^1, N_k^2 be any two basis matrices of $\mathcal{N}_r^k$. Then, there exists a square and invertible H matrix such that $N_k^2 = N_k^1 H$. Let $\mathcal{P}_k[(s,1);N_k^i] = \{\underline{x}_1^i(s),\ldots,\underline{x}_{\theta_k}^i(s)\}$, i=1,2 be the vectors generating the $\mathbb{R}[s]$-modules $\mathcal{M}[N_k^1], \mathcal{M}[N_k^2]$ respectively. Following identical steps as in the Proof of Theorem (5.4), it follows that $N_k^2 = N_k^1 H$ implies that

$$[\underline{x}_1^2(s),\ldots,\underline{x}_{\theta_k}^2(s)] = [\underline{x}_1^1(s),\ldots,\underline{x}_{\theta_k}^1(s)]H \tag{5.65a}$$

Note that $\mathcal{P}_k[(s,1);N_k^i]$ generates $\mathcal{M}[N_k^i]$, i=1,2; given that H is invertible, (5.65a) implies that $\mathcal{M}[N_k^1] = \mathcal{M}[N_k^2]$ and thus any $\mathcal{M}[N_k]$ is an invariant of $\mathcal{N}_r^k$. The invariance of $\mathcal{M}[N_k]$ suggests that for the study of its properties we may select any basis of $\mathcal{N}_r^k$. Let us consider as a basis matrix for $\mathcal{N}_r^k$ the k-th element of a complete family, say N_k^*. Such a basis matrix is expressed as in (5.64b) and thus, by Lemma (5.5) it follows that $\mathcal{M}[N_k^*]$, and thus $\mathcal{M}[N_k]$, is generated by the vectors

$$\mathcal{P}_\Omega = \{\mathcal{P}_{\sigma_1}[(s,1);\tilde{N}_{\sigma_1}];\ldots;\mathcal{P}_{\sigma_\nu}[(s,1);\tilde{N}_{\sigma_\nu}]\} \tag{5.65b}$$

By the completeness property of $\mathcal{P}(A_{\mathcal{G}})$ it follows that the completeness of $\mathcal{P}_\Omega$ and thus $\mathcal{P}_\Omega$ is a minimal basis of $\mathcal{M}[N_k]$; the existence of a minimal basis implies that $\mathcal{M}[N_k]$ is maximal $\sum_{t=1}^{\nu} \rho_i$ rank Noetherian module and since the set $\mathcal{M}[A_{\mathcal{G}}]$ is linearly independent $\mathcal{M}[N_k]$ may be expressed as in (5.62).

□

Corollary (5.8): Let $A_{\mathcal{G}} = \{\tilde{N}_{\sigma_i}, i \in \underline{\mu}\}$ be a set of generators of a complete family $\mathcal{B}(A_{\mathcal{G}})$, $\mathcal{B} = \{N_k : k \geq \sigma_1\}$ be any sequence of basis matrices of $\mathcal{S}_r(F,G)$ and let $\mathcal{W}(\mathcal{B}) = \{\mathcal{W}(N_k) : k \geq \sigma_1\}$ be the corresponding sequence of supporting subspaces.

(i) For every pair of integers (i,j) for which either $\sigma_t \leq i, j < \sigma_{t+1}$, or $\sigma_\mu \le i,j$, $\sigma_\mu = \max\{\mathcal{K}(F,G)\}$, then $\mathcal{W}(N_i)=\mathcal{W}(N_j)$.

(ii) For every pair of integers (i,j) for which $\sigma_t \le i < \sigma_{t+1} \le \sigma_p < j$, then $\mathcal{W}(N_i) \subset \mathcal{W}(N_j)$.

(iii) If σ_μ is the right index of (F,G), then $\mathcal{W}(N_{\sigma_\mu}) = \mathcal{R}^*$ is the maximal right annihilating space of (F,G) and

$$\mathcal{R}^* = \mathcal{W}(\tilde{N}_{\sigma_1}) \oplus \dots \oplus \mathcal{W}(\tilde{N}_{\sigma_\mu}) \tag{5.66}$$

□

The proof of the result is a straightforward consequence of part (ii) of Theorem (5.5).

<u>Corollary (5.9)</u>: Let $A_{\mathcal{G}}=\{\tilde{N}_{\sigma_i}, i \in \underline{\mu}\}$ be a set of generators of a complete family $\mathcal{B}(A_{\mathcal{G}})$, $\mathcal{B}=\{N_k : k \ge \sigma_1\}$ be any sequence of basis matrices of $\mathcal{S}_r(F,G)$ and let $\mathcal{M}[\mathcal{B}]=\{\mathcal{M}[N_k] : k \ge \sigma_1\}$ be the corresponding sequence of $\mathbb{R}[s]$-modules.

(i) For every pair of integers (i,j) for which either $\sigma_t \le i,j < \sigma_{t+1}$, or $\sigma_\mu \le i,j$, $\sigma_\mu = \max\{\mathcal{K}(F,G)\}$, then $\mathcal{M}[N_i]=\mathcal{M}[N_j]$.

(ii) For every pair of integers (i,j) for which $\sigma_t \le i < \sigma_{t+1} \le \sigma_p < j$, then $\text{rank}\{\mathcal{M}[N_i]\} < \text{rank}\{\mathcal{M}[N_j]\}$ and $\mathcal{M}[N_i] \subset \mathcal{M}[N_j]$.

(iii) If σ_μ is the right index of (F,G), then $\mathcal{M}[N_{\sigma_\mu}]=\mathcal{M}^*$ is the maximal $\mathbb{R}[s]$-Noetherian module which is contained in $\mathcal{N}_r(sF-G)$.

(iv) For every set of generators $A_{\mathcal{G}}$, the set of polynomial vectors

$$\mathcal{P}[s;A_{\mathcal{G}}]=\{\mathcal{P}_{\sigma_1}[(s,1);\tilde{N}_{\sigma_1}];\dots;\mathcal{P}_{\sigma_\mu}[(s,1);\tilde{N}_{\sigma_\mu}]\} \tag{5.67}$$

defines a minimal basis for $\mathcal{M}^*$ and thus for $\mathcal{N}_r(sF-G)$ with a set of minimal indices $\mathcal{I}_c(F,G)=\{(\sigma_i-1,\rho_i), i \in \underline{\mu}\}$.

<u>Proof</u>

Parts (i), (ii) and (iii) follow immediately from part (iv) of Theorem (5.5). The completeness of $\mathcal{P}(A_{\mathcal{G}})$, implies that the set

$P[s;A_G]$ is independent, has no zeros and it is column reduced. Given that $P[s;A_G]$ is a basis for M^*, it follows that it is a minimal basis and its degree set is $I_c(F,G)=\{(\sigma_i-1,\rho_i), i\in\underset{\sim}{\mu}\}$. □

The results stated above for the $M[B]$ sequence of $\mathbb{R}[s]$-modules may be interpreted in an obvious manner for the $\hat{M}[B]=\{\hat{M}[N_k];k\geq\sigma_1\}$ sequence of $\mathbb{R}[\hat{s}]$-modules. Clearly, Corollary (5.9) provides alternative means for the construction of minimal bases which are independent from the algebraic tools, used for their definition. Similarly, Corollary (5.8) provides a procedure for computing R^*, which is also independent from the construction of a minimal basis. The key tools in both procedures is the construction of a set A_G of generators of a complete family of basis matrices of $S_r(F,G)$.

<u>Remark (5.13)</u>: For any set of generators A_G, the set of homogeneous binary polynomials $P(A_G)=\{P_{\sigma_i}[(s,\hat{s});\tilde{N}_{\sigma_i}], i\in\underset{\sim}{\mu}\}$ is a homogeneous minimal basis for $N_r(sF-\hat{s}G)$. □

The last problem considered in this section is that of determining the possible degrees of prime right annihilating vectors; such a problem is intimately related to the study of possible dimensions of right annihilating spaces. In the context of linear systems, this problem has been first studied by Warren and Eckberg [War & Eck - 1] in their investigation of the possible dimensions of controllability subspaces. Using the properties of the canonical basis matrix or right annihilating vectors of an entirely right singular pencil, an alternative proof to the Warren and Eckberg result was given in [Kar - 2]. The proof of the result stated next is a straightforward generalisation of the proof given in [Kar - 2] for entirely right singular pencils.

<u>Theorem (5.6)</u>: Let $sF-\hat{s}G \in \mathbb{R}^{m\times m}[s,\hat{s}]$ be a right singular pencil and let $I_c(F,G)=\{(\varepsilon_i,\rho_i), i\in\underline{\mu}, 0\leq\varepsilon_1<\varepsilon_2<\ldots<\varepsilon_\mu\}$ be the set of c.m.i. Let us assume that $\underline{x}(s,\hat{s})=X_{d-1}\underline{e}_{d-1}(s,\hat{s})$ be a prime right annihilating vector and that ε_ν is the maximal c.m.i. for which $\varepsilon_\nu+1\leq d$. Then,

(i) If $\varepsilon_\nu+1\leq d\leq \sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)$, there always exists at least one prime vector $\underline{x}(s,\hat{s})$ with degree d-1.

(ii) If $d>\sum_{i=1}^{\nu}\rho_i(\varepsilon_i+1)$, there exists no prime vector $\underline{x}(s,\hat{s})$ with degree d-1.

<u>Proof</u>

If $\underline{x}(s,\hat{s})=X_{k-1}\underline{e}_{k-1}(s,\hat{s}) \in N_r\{sF-\hat{s}G\}$, then by Proposition (5.9)

$$X_{k-1} = Q\left[\begin{array}{c} D_{k,o,g} \\ \hdashline D_{k,\varepsilon_{g+1}} \\ \hdashline \vdots \\ \hdashline D_{k,\varepsilon_p} \\ \hdashline 0 \end{array}\right] = QD_k^c \tag{5.67a}$$

where $\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}=I_c(F,G)$ and $D_{k,\varepsilon_i}\in\mathbb{R}^{(\varepsilon_i+1)\times k}$ is the canonical block (5.18) if $k\geq\varepsilon_i+1$ and $D_{k,\varepsilon_i}=0_{\varepsilon_i+1,k}$ if $k<\varepsilon_i+1$. If we partition Q into column blocks according to the partitioning of D_k^c i.e. $Q=[Q_g,Q_{\varepsilon_{g+1}},\ldots,Q_{\varepsilon_p},\tilde{Q}]$, $Q_{\varepsilon_i}\in\mathbb{R}^{n\times(\varepsilon_i+1)}$, then (5.67a) yields

$$X_{k-1} = \left[Q_g,Q_{\varepsilon_{g+1}},\ldots,Q_{\varepsilon_p}\right]\left[\begin{array}{c} D_{k,o,g} \\ \hdashline D_{k,\varepsilon_{g+1}} \\ \hdashline \vdots \\ \hdashline D_{k,\varepsilon_p} \end{array}\right] = Q_\varepsilon\tilde{D}_k^c \tag{5.67b}$$

Clearly, $N_r\{Q_\varepsilon\}=\{0\}$ and thus the linear dependence, or independence properties of the columns of X_{k-1} is defined by the dependence, or independence properties of the columns of $\tilde{D}^c_k$. The proof of the result is then reduced to that of Theorem (5) [Kar - 2]. □

Theorem (5.5) and its Corollaries have established the existence, as well as the properties of the families of right annihilating spaces $\mathcal{W}(\mathcal{B})$ and right annihilating modules $\mathcal{M}[\mathcal{B}]$ ($\hat{\mathcal{M}}[\mathcal{B}]$). The families $\mathcal{W}(\mathcal{B}),\mathcal{M}[\mathcal{B}]$ are independent from the family of basis matrices $\mathcal{B}$, which have been used for their definition, and they characterise the pencil $sF-\hat{s}G$, or the pair (F,G). $\mathcal{W}(\mathcal{B})$ will be referred to as the right family of vector spaces of (F,G) and $\mathcal{M}[\mathcal{B}]$ as the right family of $\mathbb{R}[s]$-modules of (F,G) and because they are independent from the particular family $\mathcal{B}$ will be simply denoted by $\mathcal{W}(F,G)=\{\mathcal{W}_k:k\geq\sigma_1\}$ and $\mathcal{M}[F,G]=\{\mathcal{M}_k:k\geq\sigma_1\}$ respectively. Note that both families $\mathcal{W}(F,G),\mathcal{M}[F,G]$ are partially ordered by the set $\mathcal{K}(F,G)=\{(\sigma_i,\rho_i),\ i\in\underset{\sim}{\mu},0<\sigma_1<\ldots<\sigma_\mu\}$ since for $\mathcal{W}(F,G)$ we have that

$$\mathcal{W}_i = \mathcal{W}_j \quad \forall\, i\epsilon j \in [\sigma_\nu,\sigma_{\nu+1})\ , \quad \mathcal{W}_i = \mathcal{W}_j \quad \forall\, i,j\geq\sigma_\mu \tag{5.68a}$$

and

$$\mathcal{W}_{\sigma_1}\subset\mathcal{W}_{\sigma_2}\subset\ldots\subset\mathcal{W}_{\sigma_\mu} \tag{5.68b}$$

and similarly for $\mathcal{M}[F,G]$ we have that

$$\mathcal{M}_i = \mathcal{M}_j \quad \forall\, i\epsilon j \in [\sigma_\nu,\sigma_{\nu+1})\ , \quad \mathcal{M}_i = \mathcal{M}_j \quad \forall\, i,j\geq\sigma_\mu \tag{5.69a}$$

and

$$\mathcal{M}_{\sigma_1}\subset\mathcal{M}_{\sigma_2}\subset\ldots\subset\mathcal{M}_{\sigma_\mu} \tag{5.69b}$$

By $i\in[\sigma_\nu,\sigma_{\nu+1})$ we mean that i takes values from the set of integers $\{\sigma_\nu,\sigma_\nu+1,\ldots,\sigma_{\nu+1}-1\}$. Note that

$$\dim\mathcal{W}_{\sigma_\tau} = \operatorname{rank}\mathcal{M}_{\sigma_\tau} = \sum_{j=1}^{\tau}\rho_j\sigma_j\ , \quad \forall\,\tau\in\underset{\sim}{\mu} \tag{5.70}$$

Both families $\mathcal{W}(F,G),\mathcal{M}[F,G]$ are bounded and $\sup\mathcal{W}(F,G)=\mathcal{W}_{\sigma_\mu}=\mathcal{R}^*$ and $\sup\mathcal{M}[F,G]=\mathcal{M}_{\sigma_\mu}=\mathcal{M}^*$. $\mathcal{R}^*$ is the maximal right annihilating space of (F,G) and $\mathcal{M}^*$ is the maximal $\mathbb{R}[s]$-module in $\mathcal{N}_r\{sF-G\}$.

Remark (5.14): The families $\mathcal{W}(F,G),\mathcal{M}(F,G)$ are invariants of the equivalence class $E_s^r(F,G)=\{(F',G'):F'=RF,G'=RG,R\in\mathbb{R}^{m\times m},|R|\neq 0\}$. □

The results presented in this chapter for right Toeplitz matrices and the right annihilating spaces may be readily translated to the case of left Toeplitz matrices and left annihilating spaces; this is achieved by considering the pair (F^t,G^t) and then by transposed duality interpreting the results for (F,G).

5.6 Conclusions

The results of this chapter extend the number theoretic, geometric and algebraic results of Chapter 4 for regular pencils to the case of singular pencils. The sets $\mathcal{I}_c(F,G),\mathcal{I}_r(F,G)$ of c.m.i., r.m.i. respectively of a singular pencil have been defined by the study of discontinuity properties of Piecewise Arithmetic Progression sequences defined on the pair (F,G); this definition demonstrates the unity (from the number theoretic viewpoint) between the minimal indices and the Segré characteristics. The computation of $\mathcal{I}_c(F,G),\mathcal{I}_r(F,G)$, which is based on the Piecewise Arithmetic Progression Sequence Diagram, or Weyr Sequence Diagram, or on the set $\mathcal{K}(F,G)$, is independent from the algebraic definition of those two sets (degrees of a minimal basis); the key numerical tool involved in such computations is the computation of the rank of real matrices. An additional advantage of this approach is that the computation of those two sets is also independent from the Kronecker canonical form based procedures [Van Door - 1],[Wil - 1]. Apart from the obvious computational advantages, the present approach allows the definition of the sets of c.m.i., r.m.i. as functions defined on an ordered pair (F,G) and independently from the associated pencil.

The study of properties of the Toeplitz matrices of (F,G) has revealed a number of important aspects of the families of right (left) spaces and right (left) modules of the pair (F,G). The families $\mathcal{W}(F,G)$,$\mathcal{M}[F,G]$ of subspaces and modules have been shown to be invariants of the rational vector space $\mathcal{N}_r\{sF-G\}$; both families have also been defined geometrically and independent from their algebraic nature. The study of properties of $\mathcal{W}(F,G)$,$\mathcal{M}[F,G]$ families has led to a new procedure for computing minimal bases for the rational vector spaces $\mathcal{N}_r\{sF-\hat{s}G\}$,$\mathcal{N}_\ell\{sF-\hat{s}G\}$, which once more is purely geometric and independent from the algebraic definition of minimal bases.

One of the key tools in the development of the results of this section has been the theory of naturally partitioned matrices, also developed in this chapter. These results are general and they do not depend on the case of matrix pencils; it is believed that use of these tools to the general case of polynomial matrices may lead to general results concerning the properties and the computation of minimal bases of rational vector spaces. Finally, it is worth pointing out that the results of this chapter, on the computation of $\mathcal{I}_c(F,G)$,$\mathcal{I}_r(F,G)$ sets, may be combined with those of the previous section, on the computation of the Segré characteristics, to provide a procedure for the computation of the Kronecker canonical form of a singular pencil. The key characteristic of this new procedure is the analysis of the singularities of appropriate Piecewise Arithmetic Progression sequences, which leads to the derivation of the Kronecker form without resorting to the use of special transformations.

CHAPTER 6: BILINEAR-STRICT EQUIVALENCE OF MATRIX PENCILS

CHAPTER 6:
Bilinear–Strict Equivalence of Matrix Pencils

6.1 Introduction

The study of properties of the matrix pencils $sF-G$ and $F-sG$, defined from an ordered [illegible] (F,G), has demonstrated the existence of an important notion of duality, the so-called "elementary divisor type duality", or "integrator-differentiator type duality" [Kar. & Hay. - 1,2]. The matrix pencils $sF-G$, $F-sG$ are related by the special type of bilinear transformation: $s \to 1/\lambda$, which simply transforms the points $0, \infty, \alpha \neq 0$ of the compactified complex plane $(\mathbb{C} \cup \{\infty\})$ (or of the Riemann sphere) to the points $\infty, 0, 1/\alpha$ correspondingly. The duality notion between $sF-G$ and $F-sG$ stems from the nature of this special bilinear transformation. The aim of this chapter is to study the properties of matrix pencils under the combined action of the general bilinear and strict equivalence groups acting on a homogeneous matrix pencil $sF-\lambda G$, or on the ordered pair $(F,-G)$. A complete set of invariants for matrix pencils under Bilinear-Strict Equivalence will be defined; this work forms the basis for the study of various notions of duality defined on linear systems and provides the means for the development of a unifying approach for the study of properties of regular and extended state space systems.

The study of the Bilinear-Strict Equivalence of matrix pencils has been initiated by the work of Turnbull and Aitken [Tur & Ait - 1]; in [Tur & Ait - 1] the invariance property of invariant polynomials and the invariance of the minimal indices is established. The results in this chapter complete the work in [Tur & Ait - 1] by defining a complete set of invariants for the Bilinear-Strict Equivalence class of a matrix pencil, or of an ordered pair (F,G). An essential part in the search for a complete set of invariants of $sF-\lambda G$ under Bilinear-Strict Equivalence is the study of invariants of homogeneous binary polynomials, [illegible], under Projective-Equivalence

CHAPTER 6: BILINEAR-STRICT EQUIVALENCE OF MATRIX PENCILS

6.1 Introduction

The study of properties of the matrix pencils $sF-G$ and $F-\hat{s}G$, defined from the homogeneous pencil $sF-\hat{s}G$, or defined on a pair $(F,-G)$, has demonstrated the existence of an important notion of duality, the so-called "elementary divisor type duality", or "integrator-differentiator type duality" [Kar. & Hay - 1,2]. The matrix pencils $sF-G, F-\hat{s}G$ are related by the special type of bilinear transformation: $s \to 1/\hat{s}$, which clearly transforms the points $0,\infty,\alpha\neq 0$ of the compactified complex plane $(\mathbb{C}\cup\{\infty\})$ (or of the Riemann sphere) to the points $\infty,0,1/\alpha$ correspondingly. The duality notion between $sF-G$ and $F-\hat{s}G$ stems from the nature of this special bilinear transformation. The aim of this chapter is to study the properties of matrix pencils under the combined action of the general bilinear and strict equivalence groups acting on a homogeneous matrix pencil $sF-\hat{s}G$, or on the ordered pair $(F,-G)$. A complete set of invariants for matrix pencils under Bilinear-Strict Equivalence will be defined; this work forms the basis for the study of various notions of duality defined on linear systems and provides the means for the development of a unifying approach for the study of properties of regular and extended state space systems.

The study of the Bilinear-Strict Equivalence of matrix pencils has been initiated by the work of Turnbull and Aitken [Tur & Ait - 1]; in [Tur & Ait - 1] the covariance property of invariant polynomials and the invariance of the minimal indices is established. The results in this chapter complete the work in [Tur & Ait - 1] by defining a complete set of invariants for the Bilinear-Strict Equivalence class of a matrix pencil, or of an ordered pair (F,G). An essential part in the search for a complete set of invariants of $sF-\hat{s}G$ under Bilinear-Strict Equivalence is the study of invariants of homogeneous binary polynomials, $f(s,\hat{s})$, under Projective-Equivalence transformations; such a study of invariants is reduced to an equivalent study of invariants of matrices under Extended-Hermite type transformations. The final result is that set of column, row minimal indices, the set of all index sets of (F,-G), a set of vectors defined as Plücker vectors, and a set of Grassmann-type vectors, emerge as the set of complete and independent invariants for the Bilinear-Strict Equivalence class of a given matrix pencil. These results provide the means for the study of properties of linear state space systems under space and frequency coordinate transformations. The application of this theory to linear systems is discussed in Chapter 8.

The chapter is structured as follows: The notion of Bilinear-Strict Equivalence of homogeneous matrix pencils and an interpretation of this equivalence as coordinate transformations in the frequency domain is given in Section (6.2). In Section (6.3), a number of preliminary results are summarised; the covariance of the Smith form of $sF-\hat{s}G$ and the invariance of the column and row minimal indices under Bilinear-Strict Equivalence is established, and the notion of Projective-Equivalence of binary homogeneous polynomials is introduced. Section (6.4) is devoted to the study of Projective-Equivalence (PE) of homogeneous polynomials $f(s,\hat{s})$. The complex and real lists of $f(s,\hat{s})$ are shown to be invariant under PE and the search for a complete set of invariants under PE is finally reduced to a study of invariants of matrices under Extended Hermite Equivalence (EHE). The theory of EHE of matrices is developed in Section (6.5); the set of r-prime Plücker vectors, and the (r,i_1,i_2)-$\mathbb{C}$-canonical Grassmann vector are shown to be equivalent complete invariants under complex EHE. For the case of real EHE, it is proved that the (r,i_1,i_2)-$\mathbb{R}$-canonical Grassmann vector is a complete invariant. In Section (6.6), a complete set of invariants for binary polynomials and sets of binary polynomials under PE is derived. For both cases (a single polynomial and a set of polynomials), it is shown that the complete set of invariants is made up from the real and complex lists and the (r,i_1,i_2)-$\mathbb{R}$-canonical Grassmann vectors (characterising the real Extended Hermite Equivalence). Finally, in Section (6.7) a complete set of invariants for Bilinear-Strict Equivalence of matrix pencils is derived by combining the results of Sections (6.3) and (6.6).

6.2 Definitions, statement of the problem and interpretation of bilinear transformations

The study of Bilinear-Strict Equivalence is greatly simplified by the introduction of appropriate notation and definitions. Thus, let $\mathcal{L}=\{L:L=(F,-G),F,G\in\mathbb{R}^{m\times n}\}$ be the set of ordered pairs of m×n matrices and let $\Theta=\{\theta:\theta=(\lambda,\hat{\lambda})\}$ be the set of ordered pairs of indeterminates. For $\forall\ L=(F,-G)\in\mathcal{L}$ and $\theta=(s,\hat{s})\in\Theta$, the matrix $[L]=[F,-G]\in\mathbb{R}^{m\times 2n}$ will be called a matrix representation of L and the homogeneous polynomial matrix

$$L_\theta \overset{\Delta}{=} L(s,\hat{s}) = [F,-G]\begin{bmatrix} sI_n \\ \hat{s}I_n \end{bmatrix} = sF-\hat{s}G \tag{6.1}$$

will be referred to as the <u>θ-matrix pencil of L</u>. We define the following sets of matrix pencils: $\mathcal{L}_\theta \overset{\Delta}{=} \{L_\theta$: for a fixed $\theta=(s,\hat{s})\in\Theta$ and $\forall\, L=(F,-G)\in\mathcal{L}\}$, $\mathcal{L}(\Theta) \overset{\Delta}{=} \{L_\theta: \forall\ \theta=(s,\hat{s})\in\Theta$ and $\forall\ L=(F,-G)\in\mathcal{L}\}$. In the following, three types of equivalence are defined on $\mathcal{L}$, or equivalently on $\mathcal{L}(\Theta)$. These equivalence relations are generated by the action of appropriate transformation groups acting on $\mathcal{L}$, or $\mathcal{L}(\Theta)$.

Consider first the set $\mathcal{K}=\{k:k=(M,N),M\in\mathbb{R}^{m\times m}, N\in\mathbb{R}^{2n\times 2n}, |M|,|N|\neq 0\}$ and a composition rule (*) defined on $\mathcal{K}$ as follows: $*:\mathcal{K}\times\mathcal{K}\to\mathcal{K}$: $\forall\ k_1=(M_1,N_1)$, $k_2=(M_2,N_2)\in\mathcal{K}$, then

$$k_1 * k_2 \overset{\Delta}{=} (M_1,N_1)*(M_2,N_2) = (M_1M_2,N_2N_1) \tag{6.2}$$

It may be readily verified that $(\mathcal{K},*)$ is a group with identity element

(I_m,I_{2n}). The action of K on L is defined by: $\circ : K\times L\rightarrow L$: $\forall\, k=(M,N)\in K$ and for an $L=(F,-G)\in L$, then

$$k\circ L \overset{\Delta}{=} k\circ(F,-G) = L' = (F',-G')\in L\text{: } [L'] = M[L]N \qquad (6.3a)$$

or equivalently

$$[F',-G'] = M[F,-G]N \qquad (6.3b)$$

Such an action defines an equivalence relation E_K on L, and $E_K(L)$ denotes the equivalence class, or orbit, of $L\in L$ under K. Three important subgroups of K, and thus three notions of equivalence on L or $L(\Theta)$ are defined next:

(i) STRICT EQUIVALENCE: The subgroup $(H,*)$ of $(K,*)$, where

$$H = \{\, h\text{: } h=(R,P), R\in\mathbb{R}^{n\times n}, P=\text{diag}\{Q,Q\}, Q\in\mathbb{R}^{n\times n}, |R|,|Q|\neq 0\} \qquad (6.4a)$$

is called the Strict-Equivalence Group (SEG). The action of H on L is defined by $\circ : H\times L\rightarrow L$: $\forall\, h\in H$ and for an $L=(F,-G)\in L$, then

$$h\circ L \overset{\Delta}{=} (R,P)\circ(F,-G) = L' = (F',-G')\in L\text{: } [L'] = R[F,-G]P \qquad (6.4b)$$

or equivalently:

$$[F',-G'] = R[F,-G]\begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = [RFQ,-RGQ] \qquad (6.4c)$$

The equivalence relation E_H defined on L as above, is called Strict-Equivalence (SE) ([Gan - 1]). Two pencils $L_\theta^1=sF_1-\hat{s}G_1$, $L_\theta^2=sF_2-\hat{s}G_2\in L_\theta$ are said to be strict equivalent, $L_\theta^1 E_H L_\theta^2$, iff there exist $h\in H$: $(F_2,-G_2)= =h\circ(F_1,-G_1)$. By $E_H(F,G)$ we denote the SE class or orbit of $L_\theta=sF-\hat{s}G$, or equivalently of $L=(F,-G)$.

(ii) BILINEAR EQUIVALENCE: The subgroup $(B,*)$ of $(K,*)$, where

$$B = \{b\text{: } b=(I_m, \begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix}) = (I_m,T_\beta), \beta=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\in\mathbb{R}^{2\times 2}, |\beta|=ad-bc\neq 0\} \qquad (6.5a)$$

is called the Bilinear-Equivalence Group (BEG) [Kar & Kal - 1]. Every $b \in \mathcal{B}$ is generated by a projective transformation β, the meaning of which will be discussed later. The action of $\mathcal{B}$ on L is defined by:

$$b \circ L \triangleq b \circ (F,-G) = L' = (F',-G') \in L \colon [L'] = I_m [L] T_\beta \qquad (6.5b)$$

or equivalently:

$$[F',-G'] = I_m [F,-G] \begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix} = [aF-cG, bF-dG] \qquad (6.5c)$$

The equivalence relation $E_{\mathcal{B}}$, defined on L, is called Bilinear-Equivalence (BE). Two pencils $L^1_\theta = sF_1 - \hat{s}G_1, L^2_{\theta'} = \lambda F_2 - \hat{\lambda} G_2, \theta = (s,\hat{s}), \theta' = (\lambda, \hat{\lambda}) \in \Theta$, are said to be bilinearly equivalent, $L^1_\theta E_{\mathcal{B}} L^2_{\theta'}$, iff there exists a transformation $\beta \colon (s,\hat{s}) \to (\lambda,\hat{\lambda})$ and thus a $b \in \mathcal{B}$, generated by β, such that: $(F_2,-G_2) = b \circ (F_1,-G_1)$. By $E_{\mathcal{B}}(F,G)$ we denote the BE class of $L_\theta = sF - \hat{s}G$, or equivalently of $L = (F,-G)$.

Note that the composition rule (*) is not commutative on K; however, the following property holds true:

Proposition (6.1): For $\forall\ b \in \mathcal{B}$ and $\forall\ h \in H$, then $b * h = h * b$.

Proof

$$h * b = \left(R, \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} \right) * \left(I_m, \begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix} \right) = \left(RI_m, \begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} \right)$$

$$= \left(I_m R, \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix} \right) = b * h$$

□

Remark (6.1): $b * h \in K$, but generally $b * h \notin \mathcal{B}$ and $b * h \notin H$.

(iii) BILINEAR-STRICT EQUIVALENCE: The subgroup $(H\text{-}\mathcal{B}, *)$ of $(K, *)$, where

$$H\text{-}\mathcal{B} = \{ r \colon r = h * b,\ \forall\ h \in H,\ \forall\ b \in \mathcal{B} \} \qquad (6.6a)$$

is called the Bilinear-Strict Equivalence Group (BSEG). The action of $H\text{-}\mathcal{B}$ on L is defined by $\circ \colon H\text{-}\mathcal{B} \times L \to L \colon \forall\ r \in H\text{-}\mathcal{B}$ and for an $L = (F,-G) \in L$, then

$$r \circ L \overset{\Delta}{=} (h * b) \circ (F,-G) = h \circ \{ b \circ (F,-G)\} =$$

$$= b \circ \{ h \circ (F,-G)\} = (aRFQ-cRGQ, bRFQ-dRGQ) \tag{6.6b}$$

The equivalence relation E_{H-B} defined on L is called <u>Bilinear-Strict Equivalence</u> (BSE) [Kar & Kal - 1]. Two pencils $L_\theta^1 = sF_1 - \hat{s}G_1, L_{\theta'}^2 = \lambda F_2 - \hat{\lambda} G_2 \in L(\Theta)$, $\theta = (s,\hat{s}), \theta' = (\lambda, \hat{\lambda})$ are said to be <u>bilinearly-strict equivalent</u>, $L_\theta^1 E_{H-B} L_{\theta'}^2$, iff there exists a transformation $\beta : (s,\hat{s}) \rightarrow (\lambda,\hat{\lambda})$ and thus a $b \in B$, as well as an $h \in H$: $(F_2,-G_2) = (h * b) \circ (F_1,-G_1)$. In matrix form the above condition may be expressed by $[L_{\theta'}^2] = [(h * b) \circ L_\theta^1]$, or

$$[F_2,-G_2] = R[F_1,-G_1] \begin{bmatrix} aQ & bQ \\ cQ & dQ \end{bmatrix} \tag{6.6c}$$

By $E_{H-B}(F,G)$ we shall denote the BSE class of $L_\theta = sF - \hat{s}G$, or equivalently of $L = (F-G)$.

A preliminary result characterising the effects of $H, B, H-B$ on an $L \in L$ is stated next.

<u>Proposition (6.2)</u>: Let $L \in L, b \in B, h \in H$.

(i) If $L \xrightarrow{b} b \circ L = L^b$, then $E_H(L) \xrightarrow{b} E_H(L^b)$ is a bijection.

(ii) If $L \xrightarrow{h} h \circ L = L^h$, then $E_B(L) \xrightarrow{h} E_B(L^h)$ is a bijection.

<u>Proof</u>

(i) Let $L' \in E_H(L^b)$, then $\exists\, h \in H : L' = h \circ (L^b)$ and since $L^b = b \circ L$, we have that $L' = h \circ (b \circ L) = b \circ (h \circ L) = b \circ (L^h)$, where $L^h = h \circ L \in E_H(L)$, so $b \circ (L^h) = L'$.

If we assume that $\exists\, L^{h'} \in E_H(L)$, such that $b \circ (L^{h'}) = L'$, then $b \circ (h' \circ L) = L' = h \circ (b \circ L)$, or $h' \circ (L^b) = h \circ (L^b)$. Let us assume that $h = (R,P), h' = (R',P')$, then $R[L^b]P = R'[L^b]P'$ implies that $(R'^{-1}R, PP'^{-1}) = (I_m, I_{2n})$ and thus $R' = R, P' = P$, or $h = h'$. The proof of part (ii) follows along similar steps. □

Note that the action of $h = (R,P) \in H$ on the pencil $L_\theta = sF - \hat{s}G$ may be interpreted as

$$b\circ L_{\theta} = R[F,-G]\begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} sI_n \\ \hat{s}I_n \end{bmatrix} = R(sF-\hat{s}G)Q = sF'-\hat{s}G' = L'_{\theta} \tag{6.7}$$

and thus SE implies a coordinate transformation in the domain and codomain of the ordered pair (F,-G), but not a change in the indeterminates $(s,\hat{s})$. The action of $b=(I_m,T_\beta)$ on $L_\theta=sF-\hat{s}G$, however, may be interpreted as

$$b\circ L(s,\hat{s}) = I_m[F,-G]\begin{bmatrix} aI_n & bI_n \\ cI_n & dI_n \end{bmatrix}\begin{bmatrix} \lambda I_n \\ \hat{\lambda}I_n \end{bmatrix} =$$

$$= \lambda(aF-cG)-\hat{\lambda}(-bF+dG) = \lambda F'-\hat{\lambda}G' = L'_{\theta}, \tag{6.8}$$

which clearly expresses a change in the indeterminates from $(s,\hat{s}) \xrightarrow{\beta} (\lambda,\hat{\lambda})$ by the bilinear transformation:

$$\beta: \begin{bmatrix} s \\ \hat{s} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \lambda \\ \hat{\lambda} \end{bmatrix}, \text{ where } a,b,c,d\in\mathbb{R} \text{ and } \delta=ad-bc\neq 0 \tag{6.9}$$

Thus, a bilinear transformation expresses a coordinate type transformation in the indeterminates but not in the domain and codomain of the pair (F,-G). The action of $r=h*b$ on a matrix pencil $L_\theta=sF-\hat{s}G$ has the features of both $\mathcal{H}$ and $\mathcal{B}$ groups and thus implies a coordinate transformation in the domain and codomain of (F,-G) and a change of indeterminates from $(s,\hat{s})$ to $(\lambda,\hat{\lambda})$ according to eqn(6.9). The nature of the transformation β, that generates the transformation $b\in\mathcal{B}$ is discussed next.

It is known [Se & Kn - 1] that an n-dimensional projective domain over a field F, or a projective space $\mathbb{P}_n(F)$, is a set of entities (usually called the points of the space) that admits of a certain class $\{R\}$ of representations by homogeneous coordinates $(x_o,x_1,\ldots,x_n)$ in F; this class is such that, if R_o is any representation, the whole class $\{R\}$ consists of all those representations that can be obtained from R_o by nonsingular linear transformation

$$x'_i = \sum_{k=0}^{n} \alpha_{ik}x_k, \; i=0,1,\ldots,n$$

Thus, the representations R of $\mathbb{P}_n(F)$ are connected by a group of nonsingular linear transformations. This group is referred to as the <u>general projective group</u> and it is denoted by $PGL(n;F)$. In our case, $n=1, F=\mathbb{C}$, the projective domain $\mathbb{P}_1(\mathbb{C})$ is the <u>projective straight line</u> on the compactified complex plane $(\mathbb{C}\cup\{\infty\})$; the class $\{R\}$, is the class of all bilinear transformations $\beta:(s,\hat{s})\to(\lambda,\hat{\lambda})$ defined by

$$\beta:\ s=a\lambda+b\hat{\lambda},\ \ \hat{s}=c\lambda+d\hat{\lambda},\ \ a,b,c,d\in\mathbb{C},\ \ ad-bc\neq 0 \tag{6.10}$$

Of particular interest are the subgroups of $\{R\}$, the $\{R_{\mathbb{R}}\}$ for which $a,b,c,d\in\mathbb{R}$. The nature of the homogeneous coordinates of points in a line and the geometric meaning of β is discussed next.

On a straight line of $\mathbb{C}\cup\{\infty\}$ (Figure (6.1)), we employ two fixed points A and B as points of reference. The homogeneous coordinates of a general point P of the line, is any pair $(s,\hat{s})$ such that $s/\hat{s}=c\ \overline{AP}\,/\,\overline{PB}$, where c is a nonzero constant, and the same for all points P, while $\overline{AP},\overline{PB}$ are directed line segments, so that $\overline{AP}=-\overline{PA}$ and $\overline{PB}=-\overline{BP}$. We agree to take $\hat{s}=0$, if P=B and s=0 if P=A.

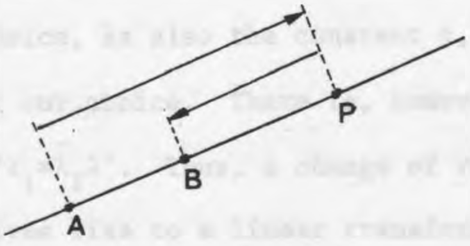

Figure (6.1)

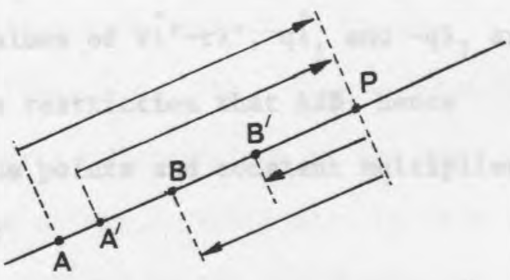

Figure (6.2)

Given P we have the ratio $s/\hat{s}$. Conversely, given the latter ratio, we have the ratio $\overline{AP}/\overline{PB}$, as well as their sum $\overline{AP}+\overline{PB}=\overline{AB}$, and hence we can find $\overline{AP}$ and therefore locate the point P.

Let us now assume that we wish to express the values of s and $\hat{s}$ in terms of the coordinates λ and $\hat{\lambda}$ of the same point P referred to new fixed points of reference A',B' (see Figure (6.2)). By definition, there is a certain new constant $k\neq 0$, such that $\lambda/\hat{\lambda}=k.\overline{A'P}/\overline{PB'}$. Since $\overline{A'P}+\overline{PB'}=\overline{A'B'}$, we may replace $\overline{A'P}$ by $\overline{A'B'}-\overline{PB'}$ and get

$$\overline{PB'} = k\hat{\lambda}.\overline{A'B'}/(\lambda+k\hat{\lambda}) \tag{6.11a}$$

Let A have the coordinates $\lambda',\hat{\lambda}'$ referred to A',B'. Then

$$\overline{PA} = \overline{PB'}-\overline{AB'} = \frac{k\hat{\lambda}.\overline{A'B'}}{\lambda+k\hat{\lambda}} - \frac{k\hat{\lambda}'.\overline{A'B'}}{\lambda'+k\hat{\lambda}'} = \frac{(\hat{\lambda}\lambda'-\lambda\hat{\lambda}')k\overline{A'B'}}{(\lambda+k\hat{\lambda})(\lambda'+k\hat{\lambda}')} \tag{6.11b}$$

Similarly, if B has the coordinates $\lambda_1,\hat{\lambda}_1$ referred to A',B', then

$$\overline{PB} = \frac{(\hat{\lambda}\lambda_1-\lambda\hat{\lambda}_1).\overline{A'B'}}{(\lambda+k\hat{\lambda})(\lambda_1+k\hat{\lambda}_1)} \tag{6.11c}$$

Hence, by dividing (6.11b),(6.11c) we have

$$\frac{s}{\hat{s}} = \frac{\overline{AP}}{\overline{PB}} = \frac{r}{q}.\frac{(\hat{\lambda}\lambda'-\lambda\hat{\lambda}')}{(\hat{\lambda}\lambda_1-\lambda\hat{\lambda}_1)}, \text{ where } \frac{r}{q} = \frac{-c(\lambda_1+k\hat{\lambda}_1)}{\lambda'+k\hat{\lambda}'} \tag{6.11d}$$

Since we are concerned only with the ratio $s/\hat{s}$, we may set

$$s = -r\hat{\lambda}'\lambda+r\lambda'\hat{\lambda},\ \hat{s} = -q\hat{\lambda}_1\lambda+q\lambda_1\hat{\lambda} \tag{6.11e}$$

Since the location of A and B with reference to A' and B' is at our choice, as also the constant c, the values of $r\hat{\lambda}'-r\lambda'$, $q\hat{\lambda}_1$ and $-q\lambda_1$ are at our choice. There is, however, the restriction that $A\not\equiv B$, hence $\hat{\lambda}'\lambda_1\neq\hat{\lambda}_1\lambda'$. Thus, a change of reference points and constant multiplier c, gives rise to a linear transformation:

$$\beta\text{: } s = a\lambda+b\hat{\lambda},\ \hat{s} = c\lambda+d\hat{\lambda},\ \delta = ad-bc \neq 0 \tag{6.11f}$$

which is clearly a coordinate transformation on the straight line. Conversely, every transformation β can be interpreted as the formulae for a change of reference points and constant multiplier.

The general projective group on the projective straight line of $\mathbb{C}\cup\{\infty\}$ will be denoted by PGL(1,ℂ); it is made up from the coordinate transformations β with $a,b,c,d\in\mathbb{C}$ and $\delta=ad-bc\neq 0$. The subgroup of PLG(1,ℂ), defined by the extra condition that $a,b,c,d\in\mathbb{R}$ will be referred to as the $\mathbb{R}$-general projective group on the projective straight line of $\mathbb{C}\cup\{\infty\}$ and shall be denoted by $PGL(1,\mathbb{C}/\mathbb{R})$. The group $PGL(1,\mathbb{C}/\mathbb{R})$ generates $\mathcal{B}$ and thus plays a crucial role in our study here.

Throughout this chapter the following notation will be adopted for the points of $\mathbb{C}\cup\{\infty\}$, or equivalently of the Riemann sphere. The equivalence class of ordered pairs $(\zeta\gamma,\zeta\delta)$, where $\gamma,\delta\in\mathbb{C}$ given $(\gamma,\delta)\neq(0,0)$ and $\zeta\in\mathbb{C}-\{0\}$, otherwise arbitrary, defines a point of $\mathbb{C}\cup\{\infty\}$, or equivalently a point on the unit radius Riemann sphere sph(Rie). The pairs $(0,1),(1,0)$ characterise the origin of $\mathbb{C}$ (the south pole of sph(Rie)), the point at infinity of $\mathbb{C}\cup\{\infty\}$ (the north pole of sph(Rie)) respectively. For every class of pairs $(\zeta\gamma,\zeta\delta)$ the invariant ratios $\zeta\gamma/\zeta\delta=\gamma/\delta$, if $\delta\neq0$ and $\zeta\delta/\zeta\gamma=\delta/\gamma$, if $\gamma\neq0$, are defined as the <u>frequency</u> and the <u>inverse frequency</u> of the class correspondingly. The inverse frequency of $(0,1)$ is defined as ∞ and the frequency of $(1,0)$ also as ∞.

By defining the reference points $(0,1)$ and $(1,0)$ on sph(Rie), every class generated by (γ,δ), with the only exception the pair $(0,0)$, uniquely defines a point on sph(Rie); such points are characterised by the invariants, the frequency and inverse frequency. The action of $\beta\in PGL(1,\mathbb{C})$ on a pair (γ,δ) may thus be interpreted as redescription of the same point with respect to a pair of new reference points on sph(Rie). In fact the parameters (a,b,c,d) of β are uniquely defined by the coordinates of the new reference points (a,b) and (c,d) with respect to the old reference points $(1,0)$ and $0,1)$ respectively.

The essence of the BSE notion, defined on matrix pencils (or equivalently on ordered pairs $(F,-G)$), is that it represents coordinate transformations on the domain and codomain of $(F,-G)$ (space coordinate transformations), as well as coordinate transformations on the sph(Rie) (frequency coordinate transformations). The study of the invariants of BSE is the main objective of this chapter; such a study will be carried out on the sets L, or equivalently on $L(\Theta)$, since transformations on L may be interpreted on $L(\Theta)$ and vice-versa. The starting point in this study is the investigation of the properties of the set of SE invariants of a matrix pencil under BE transformations. The results discussed in the following section are based on the work of Turnbull and Aitken [Tur & Ait - 1] and provide the basis for the derivation of a complete set of invariants of matrix pencils under BSE.

6.3 The properties of the set of SE invariants under BE

Let $L=(F,-G)\in L$, $\theta=(s,\hat{s})\in\Theta$ and let $L_\theta=L(s,\hat{s})=sF-\hat{s}G$ be the associated matrix pencil in $(s,\hat{s})$; $L(s,\hat{s})\in\mathbb{R}[s,\hat{s}]^{m\times n}$ and assume that $\text{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}=$ $=\rho\leq\min(m,n)$. The Smith form of $L(s,\hat{s})$ over $\mathbb{R}[s,\hat{s}]$ ([Gan - 1]) is defined by

$$S(s,\hat{s})=\begin{bmatrix} S_\rho^*(s,\hat{s}) & 0_{\rho,n-\rho} \\ 0_{m-\rho,\rho} & 0_{m-\rho,n-\rho} \end{bmatrix} \qquad (6.12a)$$

where $S_\rho^*(s,\hat{s})=\text{diag}\{\tilde{f}_1(s,\hat{s}),\tilde{f}_2(s,\hat{s}),\ldots,\tilde{f}_\rho(s,\hat{s})\}$, the $\tilde{f}_i(s,\hat{s})\in\mathbb{R}[s,\hat{s}]$ being the invariant polynomials of $L(s,\hat{s})$ over $\mathbb{R}[s,\hat{s}]$, and with the property that $\tilde{f}_i(s,\hat{s})$ divides $\tilde{f}_{i+1}(s,\hat{s})$ for $\forall\ i\in\underset{\sim}{\rho}$, with $\tilde{f}_i(s,\hat{s})=0$ for $i>\rho$. The set of $\{\tilde{f}_i(s,\hat{s}),i\in\underset{\sim}{\rho}\}$ is defined by the standard Smith algorithm ([Gan - 1]) by

$$\tilde{f}_i(s,\hat{s})\triangleq d_i(s,\hat{s})/d_{i-1}(s,\hat{s}),\ i\in\underset{\sim}{\rho}\text{ and } d_o(s,\hat{s})=1 \qquad (6.12b)$$

where $d_i(s,\hat{s})$ is the ith-determinantal divisor of $L(s,\hat{s})$, that is the g.c.d. of all $i\times i$ minors of $L(s,\hat{s})$. If there are k nonzero trivial elements in $\{\tilde{f}_i(s,\hat{s}),i\in\underset{\sim}{\rho}\}$, i.e. $S_\rho^*(s,\hat{s})=\text{diag}\{\underbrace{1,\ldots,1}_{k},f_1(s,\hat{s}),\ldots,f_{\rho-k}(s,\hat{s})\}$, then k will be referred to as the <u>power</u> of $L(s,\hat{s})$ and the ordered set of homogeneous invariant polynomials $F(F,G)=\{f_i(s,\hat{s}),i\in\underset{\sim}{\rho-k}$ (ordering is defined by divisibility) will be referred to as the <u>homogeneous invariant polynomial set</u> of $L(s,\hat{s})$.

It is a well known result ([Gan-1],[Tur & Ait-1]) that if $I_c(F,G)$, $I_r(F,G)$ are the c.m.i., r.m.i. respectively of $L(s,\hat{s})$, then the sets $\mathcal{F}(F,G)$ (equivalently the set of homogeneous e.d.), $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$ form a complete and independent set of invariants for the $\mathcal{E}_{\mathcal{H}}(F,G)$ equivalence class. The study of the effect of $b\in\mathcal{B}$ on the set $\{\mathcal{F}(F,G),\mathcal{I}_c(F,G),\mathcal{I}_r(F,G)\}$ of SE invariants is examined next. It will be shown that such a study is reduced to an investigation of the effects of the projective transformation $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ on homogeneous polynomials. Some preliminary definitions are given first.

Let $\mathbb{R}_d\{\Theta\}$ be the set of homogeneous polynomials of degree d with coefficients from $\mathbb{R}$, in all possible indeterminates $\theta=(s,\hat{s})\in\Theta$. Let $\beta:(s,\hat{s})\to(\lambda,\hat{\lambda})$ be the projective transformation defined by eqn(6.9). The action of β on $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$ may be defined by

$$\beta\circ f(s,\hat{s}) = \tilde{f}(\lambda,\hat{\lambda}) = f(a\lambda+b\hat{\lambda},c\lambda+d\hat{\lambda}) \tag{6.13}$$

Two polynomials $f_1(s,\hat{s}),f_2(\lambda,\hat{\lambda})\in\mathbb{R}_d\{\Theta\}$ will be said to be <u>projectively equivalent</u> (PE), and shall be denoted by $f_1(s,\hat{s})\mathcal{E}_P f_2(\lambda,\hat{\lambda})$, if there exists a $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ such that $\beta:(s,\hat{s})\to(\lambda,\hat{\lambda})$ and a $c\in\mathbb{R}-\{0\}$ such that

$$\beta\circ f_1(s,\hat{s}) = c.f_2(\lambda,\hat{\lambda}) \tag{6.14}$$

Clearly, (6.14) defines an equivalence relation $\mathcal{E}_P$, on $\mathbb{R}_d\{\Theta\}$, which is called <u>projective equivalence</u> (PE).

Let $\mathcal{F}_1=\{f_i(s,\hat{s})\in\mathbb{R}_{d_i}\{\Theta\},i\in\underline{\rho}\}$, $\mathcal{F}_2=\{\tilde{f}_i(\lambda,\hat{\lambda})\in\mathbb{R}_{d_i}\{\Theta\},i\in\underline{\rho}\}$ are two ordered sets of homogeneous polynomials. $\mathcal{F}_1,\mathcal{F}_2$ are said to be <u>projectively equivalent</u>, $\mathcal{F}_1\mathcal{E}_P\mathcal{F}_2$, if and only if $f_i(s,\hat{s})\mathcal{E}_P\tilde{f}_i(\lambda,\hat{\lambda})$ for $\forall$ $i\in\underline{\rho}$ and for the same transformation β. The projective equivalence class of $f(s,\hat{s})(\mathcal{F})$ shall be denoted by $\mathcal{E}_P(f)(\mathcal{E}_P(\mathcal{F}))$.

<u>Lemma (6.1)</u>: Let $f_1(s,\hat{s})\in\mathbb{R}_{d_1}\{\Theta\}$, $f_2(s,\hat{s})\in\mathbb{R}_{d_2}\{\Theta\}$ and let $g(s,\hat{s})\in\mathbb{R}_p\{\Theta\}$

be their greatest common divisor (gcd). Let β be any transformation in $PGL(1,\mathbb{C}/\mathbb{R})$: $(s,\hat{s}) \xrightarrow{\beta} (\lambda,\hat{\lambda})$ and let $\tilde{f}_1(\lambda,\hat{\lambda})=\beta\circ f_1(s,\hat{s})$, $\tilde{f}_2(\lambda,\hat{\lambda})=\beta\circ f_2(s,\hat{s})$ and $\tilde{g}(\lambda,\hat{\lambda})=\beta\circ g(s,\hat{s})$. Then $\tilde{g}(\lambda,\hat{\lambda})$ is a g.c.d. of $\tilde{f}_1(\lambda,\hat{\lambda}),\tilde{f}_2(\lambda,\hat{\lambda})$.

Proof

Since $g(s,\hat{s})$ is a g.c.d. of $f_1(s,\hat{s}),f_2(s,\hat{s})$, then $f_1(s,\hat{s})=h_1(s,\hat{s}).g(s,\hat{s})$, $f_2(s,\hat{s})=h_2(s,\hat{s}).g(s,\hat{s})$ where $h_1(s,\hat{s}),h_2(s,\hat{s})$ are homogeneous polynomials. Clearly, $\beta\circ f_1(s,\hat{s})=(\beta\circ h_1(s,\hat{s}))(\beta\circ g(s,\hat{s}))=\tilde{h}_1(\lambda,\hat{\lambda}).\tilde{g}(\lambda,\lambda)$, $\beta\circ f_2(s,\hat{s})=$ $=(\beta\circ h_2(s,\hat{s}))(\beta\circ g(s,\hat{s}))=\tilde{h}_2(\lambda,\hat{\lambda})\tilde{g}(\lambda,\lambda)$ and thus $\tilde{g}(\lambda,\hat{\lambda})|\tilde{f}_1(\lambda,\hat{\lambda}),\tilde{g}(\lambda,\hat{\lambda})|\tilde{f}_2(\lambda,\hat{\lambda})$. If $\tilde{g}'(\lambda,\hat{\lambda})=\gcd\{\tilde{f}_1(\lambda,\hat{\lambda}),\tilde{f}_2(\lambda,\hat{\lambda})\}$, then $\tilde{g}(\lambda,\hat{\lambda})|\tilde{g}'(\lambda,\hat{\lambda})$. By applying the inverse transformation $\beta^{-1}:(\lambda,\hat{\lambda}) \xrightarrow{\beta^{-1}} (s,\hat{s})$, then $\beta^{-1}\circ\tilde{g}'(\lambda,\hat{\lambda})=\bar{g}(s,\hat{s})$ and clearly $\bar{g}(s,\hat{s})|f_1(s,\hat{s}),\bar{g}(s,\hat{s})|f_2(s,\hat{s})$; thus, $\bar{g}(s,\hat{s})|g(s,\hat{s})$ and hence $g(s,\hat{s})=\bar{g}(s,\hat{s})\bar{h}(s,\hat{s})$. By applying β, we have $\beta\circ g(s,\hat{s})=(\beta\circ\bar{g}(s,\hat{s}))(\beta\circ h(s,\hat{s}))=$ $=\tilde{g}'(\lambda,\hat{\lambda})\tilde{h}(\lambda,\hat{\lambda})$ and hence $\tilde{g}'(\lambda,\hat{\lambda})|\tilde{g}(\lambda,\hat{\lambda})$. Since $\tilde{g}(\lambda,\hat{\lambda})|\tilde{g}'(\lambda,\hat{\lambda})$ and $\tilde{g}'(\lambda,\hat{\lambda})|\tilde{g}(\lambda,\hat{\lambda})$, then $\tilde{g}(\lambda,\hat{\lambda})$ and $\tilde{g}'(\lambda,\hat{\lambda})$ are associates. □

This property defines the covariance of the g.c.d. under PE transformations [Tur & Ait - 1]. An immediate consequence of this lemma is the following result [Tur & Ait - 1].

Proposition (6.3): Let $L_1(s,\hat{s})=sF_1-\hat{s}G_1$, $L_2(\lambda,\hat{\lambda})=\lambda F_2-\hat{\lambda}G_2\in\mathcal{L}(\Theta)$, $\mathcal{F}(F_1,G_1)=$ $=\{f_i(s,\hat{s}),i\in\underbrace{\rho_1-k}_1\}$, $\mathcal{F}(F_2,G_2)=\{\tilde{f}_i(\lambda,\hat{\lambda}),i\in\underbrace{\rho_2-k}_2\}$ be the corresponding homogeneous invariant polynomial sets of $L_1(s,\hat{s}),L_2(\lambda,\hat{\lambda})$, where (ρ_1,k_1), (ρ_2,k_2) are the respective pairs of rank, power. If $L_1(s,\hat{s})\mathcal{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$ for some $h\in\mathcal{H}$, and some $b\in\mathcal{B}$, generated by $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ such that $(s,\hat{s}) \xrightarrow{\beta} (\lambda,\hat{\lambda})$, then

(i) $\rho_1=\rho_2=\rho$ and $k_1=k_2=k$.

(ii) $\mathcal{F}(F_1,G_1)\mathcal{E}_{\mathcal{P}}\mathcal{F}(F_2,G_2)$. □

A result expressing the effect of $\mathcal{H}-\mathcal{B}$ on the sets $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$ of $L(s,\hat{s})$ is given next [Tur & Ait - 1].

Proposition (6.4): Let $L_1(s,\hat{s})=sF_1-\hat{s}G_1$, $L_2(\lambda,\hat{\lambda})=\lambda F_2-\hat{\lambda}G_2\in\mathcal{L}(\Theta)$ and let $\mathcal{I}_c(F_1,G_1),\mathcal{I}_c(F_2,G_2)$ and $\mathcal{I}_r(F_1,G_1),\mathcal{I}_r(F_2,G_2)$ be the corresponding sets of c.m.i., r.m.i of $L_1(s,\hat{s}),L_2(\lambda,\hat{\lambda})$. If $L_1(s,\hat{s})\mathcal{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$, then

$$\mathcal{I}_c(F_1,G_1) = \mathcal{I}_c(F_2,G_2) \text{ and } \mathcal{I}_r(F_1,G_1) = \mathcal{I}_r(F_2,G_2)$$

Proof

Let $U(s,\hat{s})$ be a homogeneous minimal basis for $\mathcal{N}_r\{sF-\hat{s}G\}$ and let $\underline{u}(s,\hat{s})$ be a minimal degree vector of $U(s,\hat{s})$. Note that the linear transformation $\beta:(s,\hat{s})\rightarrow(\lambda,\hat{\lambda})$ cannot raise the degree in the transformed vector $\tilde{\underline{u}}(\lambda,\hat{\lambda})=\beta\circ\underline{u}(s,\hat{s})$, which is in $\mathcal{N}_r\{\lambda\tilde{F}-\hat{\lambda}\tilde{G}\}$ ($\lambda\tilde{F}-\hat{\lambda}\tilde{G}$ is the transformed pencil); however, this transformation might lend to the lowering of the degree of $\tilde{\underline{u}}(\lambda,\hat{\lambda})$, through the cancelling of some common factor in the elements of the vector. In such a case, however, by transforming $\tilde{\underline{u}}(\lambda,\hat{\lambda})$ by $\beta^{-1}:(\lambda,\hat{\lambda})\rightarrow(s,\hat{s})$, a vector $\underline{u}'(s,\hat{s})$ in $\mathcal{N}_r\{sF-\hat{s}G\}$ is obtained, which contradicts the hypothesis that $\underline{u}(s,\hat{s})$ is a vector of the minimal basis $U(s,\hat{s})$. By the same reasoning applied to the transposed pencil $(sF-\hat{s}G)^t$, the invariance of $\mathcal{I}_r(F,G)$ is established. □

Propositions (6.3),(6.4) express respectively the covariance property of the homogeneous invariant polynomials and the invariance property of the sets of c.m.i. and r.m.i. of $L(s,\hat{s})$ under $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ equivalence. By combining the above two results, the following criterion for $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ equivalence of matrix pencils is obtained.

Theorem (6.1): Let $L_1(s,\hat{s})=sF_1-\hat{s}G_1$, $L_2(\lambda,\hat{\lambda})=\lambda F_2-\hat{\lambda}G_2\in\mathcal{L}(\Theta)$. $L_1(s,\hat{s})\mathcal{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$ if and only if the following conditions hold true:

(i) $\mathcal{I}_c(F_1,G_1)=\mathcal{I}_c(F_2,G_2)$, $\mathcal{I}_r(F_1,G_1)=\mathcal{I}_r(F_2,G_2)$.

(ii) $\mathcal{F}(F_1,G_1)\mathcal{E}_{\mathcal{P}}\mathcal{F}(F_2,G_2)$.

Proof

The necessity of the result follows immediately by combining

Propositions (6.3) and (6.4). To prove the sufficiency assume that the conditions in (i) and (ii) hold true. Since $\mathcal{F}(F_1,G_1)\mathcal{E}_{\mathcal{P}}\mathcal{F}(F_2,G_2)$, then $L_1(s,\hat{s}),L_2(\lambda,\hat{\lambda})$ have the same rank ρ $(\rho_1=\rho_2=\rho)$ and the same power k $(k_1=k_2=k)$ and there exists a transformation $\beta\epsilon PGL(1,\mathbb{C}/\mathbb{R}):(\lambda,\hat{\lambda})\xrightarrow{\beta}(s,\hat{s})$ for which $\beta\circ f_{2,i}(\lambda,\hat{\lambda})=c_{2,i}.\tilde{f}_{2,i}(s,\hat{s})\ \forall\ i\epsilon\underset{\sim}{\rho-k}$. The transformation $\beta\epsilon PGL(1/\mathbb{C}/\mathbb{R})$ generates a $b\epsilon\mathcal{B}$ and let $b\circ L_2(\lambda,\hat{\lambda})=s\tilde{F}_2-\hat{s}\tilde{G}_2=\tilde{L}_2(s,\hat{s})$. Since $\tilde{L}_2(s,\hat{s})\mathcal{E}_{\mathcal{B}}L_2(\lambda,\hat{\lambda})$, then $\mathcal{I}_c(\tilde{F}_2,\tilde{G}_2)=\mathcal{I}_c(F_2,G_2)=\mathcal{I}_c(F_1,G_1)$, $\mathcal{I}_r(\tilde{F}_2,\tilde{G}_2)=\mathcal{I}_r(F_2,G_2)=\mathcal{I}_r(F_1,G_1)$; furthermore, the sets of homogeneous invariant polynomials of $\tilde{L}_2(s,\hat{s})$ and $L_1(s,\hat{s})$ differ only by scalars (units of $\mathbb{R}[s,\hat{s}]$) (β has been constructed this way) and thus $L_1(s,\hat{s}),\tilde{L}_2(s,\hat{s})$ have the same Smith form. The pencils $\tilde{L}_2(s,\hat{s}),L_1(s,\hat{s})$ have therefore the same Smith form over $\mathbb{R}[s,\hat{s}]$ (or equivalently the same sets of e.d.) and the same sets of c.m.i. and r.m.i. and thus $\tilde{L}_2(s,\hat{s})\mathcal{E}_{\mathcal{H}}L_1(s,\hat{s})$ [Gan - 1]. Therefore, there exists an $h\epsilon\mathcal{H}$ such that $L_1(s,\hat{s})=h\circ\tilde{L}_2(s,\hat{s})$; given that $\tilde{L}_2(s,\hat{s})=b\circ L_2(\lambda,\hat{\lambda})$ it follows that $L_1(s,\hat{s})=(h*b)\circ L_2(\lambda,\hat{\lambda})$ and thus $L_1(s,\hat{s})\mathcal{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$. ☐

The key notion in the characterisation of $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ equivalence of matrix pencils is thus the notion of $\mathcal{E}_{\mathcal{P}}$ equivalence which is defined on the set of homogeneous invariant polynomials $\mathcal{F}(F,G)$ of the pencil $sF-\hat{s}G$. By definition, such a study is reduced to an investigation of the conditions under which two polynomials $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})\ \epsilon\,\mathbb{R}_d\{\Theta\}$ are $\mathcal{E}_{\mathcal{P}}$-equivalent; the problem of finding the conditions under which $f(s,\hat{s})\mathcal{E}_{\mathcal{P}}\tilde{f}(\lambda,\hat{\lambda})$ is equivalent to a problem of finding a complete and independent set of invariants for the orbit $\mathcal{E}_{\mathcal{P}}(f(s,\hat{s}))$. This problem is considered next.

6.4 Projective equivalence of homogeneous binary polynomials: preliminary results

The aim of this section is to give a number of preliminary results on the $\mathcal{E}_{\mathcal{P}}$-equivalence defined on the set $\mathbb{R}_d\{\Theta\}$, and provide the means for the definition of a complete set of invariants for the orbit $\mathcal{E}_{\mathcal{P}}(f(s,\hat{s}))$, $f(s,\hat{s})\ \epsilon\,\mathbb{R}_d\{\Theta\}$. The origins of $\mathcal{E}_{\mathcal{P}}$-equivalence stem from the classical work in [Tur - 1] on the algebraic theory of invariants; however, apart from some preliminary results on the effect of the projective transformation β on the factorisation properties of $f(s,\hat{s})$, no complete set of invariants for $\mathcal{E}_p(f)$ has been defined so far. It will be shown here that finding a complete set of invariants for $\mathcal{E}_p$-equivalence is equivalent to a classical problem of algebraic projective geometry [Se & Kn - 1] that is: find the conditions under which two symmetric sets of points of $\mathbb{C}\cup\{\infty\}$ may be transformed to each other under a projective transformation $\beta\epsilon PGL(1,\mathbb{C}/\mathbb{R})$. The latter problem will be expressed in an equivalent form as study of invariants of matrices under the notion of Extended-Hermite equivalence. We start off by investigating the effect of PE transformations on the factorisation properties of $f(s,\hat{s})\epsilon\mathbb{R}_d\{\Theta\}$.

<u>Lemma (6.2)</u> [Tur - 1]: Let $f(s,\hat{s})=rs^2+ps\hat{s}+q\hat{s}^2, \tilde{f}(\lambda,\hat{\lambda})=\tilde{r}\lambda^2+\tilde{p}\lambda\hat{\lambda}+\tilde{q}\hat{\lambda}^2\epsilon\mathbb{R}_2\{\Theta\}$ and let $\Delta=p^2-4rq, \tilde{\Delta}=\tilde{p}^2-4\tilde{r}\tilde{q}$ be their corresponding discriminants. If $f(s,\hat{s})\mathcal{E}_p\tilde{f}(\lambda,\hat{\lambda})$, i.e. $\beta\circ f(s,\hat{s})=c.\tilde{f}(\lambda,\hat{\lambda})$, then

$$\tilde{\Delta} = \delta^2/c^2.\Delta$$

where $\delta=ad-bc$ is the determinant of $\beta\epsilon PGL(1,\mathbb{C}/\mathbb{R})$.

<u>Proof</u>

We may write $f(s,\hat{s}),f(\lambda,\hat{\lambda})$ as

$$f(s,\hat{s})=[s,\hat{s}]\begin{bmatrix} r & p/2 \\ p/2 & q \end{bmatrix}\begin{bmatrix} s \\ \hat{s} \end{bmatrix}=\underline{s}^tD\underline{s},\ f(\lambda,\hat{\lambda})=[\lambda,\hat{\lambda}]\begin{bmatrix} \tilde{r} & \tilde{p}/2 \\ \tilde{p}/2 & \tilde{q} \end{bmatrix}\begin{bmatrix} \lambda \\ \hat{\lambda} \end{bmatrix}=\underline{\lambda}^t\tilde{D}\underline{\lambda} \qquad (6.15a)$$

and obviously $|D|=-1/4.\Delta$, $|\tilde{D}|=-1/4.\tilde{\Delta}$. Because $\beta:(s,\hat{s})\rightarrow(\lambda,\hat{\lambda})$ we have that

$$\begin{bmatrix} s \\ \hat{s} \end{bmatrix}=[\beta]\begin{bmatrix} \lambda \\ \hat{\lambda} \end{bmatrix}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \lambda \\ \hat{\lambda} \end{bmatrix},\ \delta=|[\beta]|=ad-bc\neq 0 \qquad (6.15b)$$

and thus

$$\beta\circ f(s,\hat{s})=([\beta]\underline{\lambda})^tD([\beta]\underline{\lambda})=\underline{\lambda}^t([\beta]^tD[\beta])\underline{\lambda}=\lambda^tD^*\underline{\lambda} \qquad (6.15c)$$

By (6.15a) and (6.15c) we have that $[\beta]^t D[\beta]=c.\tilde{D}$ and thus $\delta^2\Delta=c^2\tilde{\Delta}$.

Remark (6.1): The condition $\tilde{\Delta}=\delta^2/c^2\Delta$ implies that the discriminant Δ is an invariant of the $f(s,\hat{s})=rs^2+ps\hat{s}+q\hat{s}^2$ of weight 2 under PE transformations [Tur - 1]. Furthermore, $\mathrm{sign}\{\Delta\}=\mathrm{sign}\{\tilde{\Delta}\}$ and the reducibility properties over $\mathbb{R}$ of $f(s,\hat{s})\in\mathbb{R}_2\{\Theta\}$ are invariant under $\mathcal{E}_p$-equivalence.

By Lemma (6.2) and its remark we have the following result.

Proposition (6.5): Let $p_i(s,\hat{s})=(\gamma_i s-\delta_i\hat{s}),\gamma_i,\delta_i\in\mathbb{C}$, $(\gamma_i,\delta_i)\neq(0,0)$ be the primes over $\mathbb{C}$ of $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$. If $\tilde{f}(\lambda,\hat{\lambda})\in\mathbb{R}_d\{\Theta\}$ and $f(s,\hat{s})\mathcal{E}_p\tilde{f}(\lambda,\hat{\lambda})$, with some $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$, then:

(i) Any pair $p_i(s,\hat{s}),p_j(s,\hat{s})$ of distinct primes $((\gamma_i,\delta_i)\neq\zeta(\gamma_j,\delta_j),\zeta\in\mathbb{C}-\{0\})$ of $f(s,\hat{s})$ is mapped under β to a distinct pair of primes of $\tilde{f}(\lambda,\hat{\lambda})$ and vice-versa.

(ii) Any pair of complex conjugate primes $p(s,\hat{s})=(\gamma s-\delta\hat{s}),\bar{p}(s,\hat{s})=(\bar{\gamma}s-\bar{\delta}\hat{s})$ of $f(s,\hat{s})$ is mapped under β to a pair of complex conjugate primes of $\tilde{f}(\lambda,\hat{\lambda})$ and vice-versa.

(iii) Any pair $p(s,\hat{s}),p'(s,\hat{s})$ of repeated primes $((\gamma,\delta)=\zeta(\gamma',\delta'),\zeta\in\mathbb{C}-\{0\})$ of $f(s,\hat{s})$ is mapped under β to a pair of repeated primes of $\tilde{f}(\lambda,\hat{\lambda})$ and vice-versa. □

The proof of the above result follows immediately by the way the projective transformation $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ is applied on the unique factorisation of $f(s,\hat{s})$ and by Lemma (6.2) and its remark.

Let us now consider an $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$ and let us denote by

$$\mathcal{D}_{\mathbb{R}}(f)=\{(\alpha_i s-\beta_i\hat{s})^{\tau_i},\alpha_i,\beta_i\in\mathbb{R},(\alpha_i,\beta_i)\neq(0,0),i\in\underset{\sim}{\mu}\} \tag{6.16}$$

$$\mathcal{D}_{\mathbb{C}}(f)=\{(\gamma_i s-\delta_i\hat{s})^{p_i},(\bar{\gamma}_i s-\bar{\delta}_i\hat{s})^{p_i},\gamma_i,\delta_i\in\mathbb{C},(\gamma_i,\delta_i)\neq(0,0),i\in\underset{\sim}{\nu}\} \tag{6.17}$$

the sets of real, complex e.d. of $f(s,\hat{s})$ over $\mathbb{C}$ respectively; note that $\mathcal{D}_{\mathbb{C}}(f)$ is symmetric, i.e. if $(\gamma_i s-\delta_i\hat{s})^{p_i}\in\mathcal{D}_{\mathbb{C}}(f)$, then the complex conjugate

$(\bar{\gamma}_i s-\bar{\delta}_i \hat{s})^{p_i}\in\mathcal{D}_{\mathbb{C}}(f)$. By Proposition (6.5), the $E_{\mathcal{P}}$-equivalence may be characterised as an equivalence defined on the sets $\mathcal{D}_{\mathbb{R}}(f)$, and $\mathcal{D}_{\mathbb{C}}(f)$ in the following way:

<u>Proposition (6.6)</u>: Let $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})\in\mathbb{R}_d\{\theta\}$ and let $(\mathcal{D}_{\mathbb{R}}(f),\mathcal{D}_{\mathbb{C}}(f))$, $(\mathcal{D}_{\mathbb{R}}(\tilde{f}),\mathcal{D}_{\mathbb{C}}(\tilde{f}))$ be the corresponding unique factorisation sets defined above. $f(s,\hat{s})E_{\mathcal{P}}\tilde{f}(\lambda,\hat{\lambda})$, if and only if there exists a $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ such that the following conditions hold true:

(i) For $\forall\ e_i(s,\hat{s})\in\mathcal{D}_{\mathbb{R}}(f)$, there exists an $\tilde{e}_i(\lambda,\hat{\lambda})\in\mathcal{D}_{\mathbb{R}}(\tilde{f})$ such that $e_i(s,\hat{s})E_{\mathcal{P}}\tilde{e}_i(\lambda,\hat{\lambda})$, or equivalently: $c_i.\tilde{e}_i(\lambda,\hat{\lambda})=\beta\circ e_i(s,\hat{s}),c_i\in\mathbb{R}-\{0\}$ and vice-versa.

(ii) For $\forall\ e'_i(s,\hat{s})\in\mathcal{D}_{\mathbb{C}}(f)$, there exists an $\tilde{e}'_i(\lambda,\hat{\lambda})\in\mathcal{D}_{\mathbb{C}}(\tilde{f})$ such that $e'_i(s,\hat{s})E_{\mathcal{P}}\tilde{e}'_i(\lambda,\hat{\lambda})$, or equivalently: $c'_i.\tilde{e}'_i(\lambda,\hat{\lambda})=\beta\circ e'_i(s,\hat{s}),c'_i\in\mathbb{C}-\{0\}$ and vice-versa. □

This result expresses the covariance property of the sets $\mathcal{D}_{\mathbb{R}}(f),\mathcal{D}_{\mathbb{C}}(f)$ of $f(s,\hat{s})$ under $E_{\mathcal{P}}$-equivalence and reduces the study of invariants of $E_{\mathcal{P}}(f)$ to the study of properties of the e.d. sets under a common PE transformation β. Before we proceed to this study, we note that a real e.d. $(\alpha s-\beta\hat{s})^{\tau}$ may be represented by an ordered triple $(\alpha,\beta;\tau)$, where $\alpha,\beta\in\mathbb{R}$ and $\tau\in\mathbb{Z}$ and that a pair of complex conjugate e.d. $(\gamma s-\delta\hat{s})^{p},(\bar{\gamma}s-\bar{\delta}\hat{s})^{p}$ may be represented by an ordered triple $(\gamma,\delta;p)$, where $\gamma,\delta\in\mathbb{C}$, $p\in\mathbb{Z}$; using those two representations we may define the following sets for an $f(s,\hat{s})\in\mathbb{R}_d\{\theta\}$.

<u>Definition (6.1)</u>:

(i) We define by $\mathcal{B}'_i\triangleq\{(\gamma^i_j,\delta^i_j;p_i);\ \gamma^i_j,\delta^i_j\in\mathbb{C},p_i\in\mathbb{Z},j\in\underset{\sim}{\nu}_i,(\gamma^i_j,\delta^i_j)\neq\xi(\gamma^i_k,\delta^i_k),$ $\forall\ j\neq k,\xi\in\mathbb{C}-\{0\}\}$, as the set of all ordered triples corresponding to all pairs $(\gamma^i_j s-\delta^i_j\hat{s})^{p_i},(\bar{\gamma}^i_j s-\bar{\delta}^i_j\hat{s})^{p_i}$ in $\mathcal{D}_{\mathbb{C}}(f)$ with the same degree p_i. An ordering of the elements of $\mathcal{B}'_i$ is defined by any permutation of its elements; such a permutation will be denoted by $\pi(\mathcal{B}'_i)$ and the set of all such permutations by $<\mathcal{B}'_i>$.

(ii) $\mathcal{B}_{\mathbb{C}}(f) \overset{\Delta}{=} \{\mathcal{B}'_1;\dots;\mathcal{B}'_\sigma;p_1<\dots<p_\sigma\}$ corresponds to the set of all pairs of complex conjugate e.d. of $f(s,\hat{s})$ and shall be referred to as the complex unique factorisation set ($\mathbb{C}$-UFS) of $f(s,\hat{s})$. The set $J_{\mathbb{C}}(f) \overset{\Delta}{=} \{(p_1,\nu_1),\dots$ $\dots,(p_\sigma,\nu_\sigma)\}$, where $\nu_i=\#\mathcal{B}'_i$ (the cardinality of $\mathcal{B}'_i$), is defined as the complex list of $f(s,\hat{s})$. Every permutation of $\mathcal{B}_{\mathbb{C}}(f)$ of the type $\pi(\mathcal{B}_{\mathbb{C}}(f))=\{\pi(\mathcal{B}'_1),\dots,\pi(\mathcal{B}'_\sigma):\pi(\mathcal{B}'_i)\in<\mathcal{B}'_i>\}$ defines a normal ordering of $\mathcal{B}_{\mathbb{C}}(f)$ and the set of all such permutations will be denoted by $<\mathcal{B}_{\mathbb{C}}(f)>$.

(iii) Let $\pi(\mathcal{B}_{\mathbb{C}}(f))=\{\pi(\mathcal{B}'_1),\dots,\pi(\mathcal{B}'_\sigma)\}\in<\mathcal{B}_{\mathbb{C}}(f)>$, where $\pi(\mathcal{B}'_i)= (\gamma^i_1,\delta^i_1;p_i),\dots$ $\dots,(\gamma^i_{\nu_i},\delta^i_{\nu_i};p_i)\}\in<\mathcal{B}'_i>$. A matrix representation of $\pi(\mathcal{B}_{\mathbb{C}}(f))$ may then be defined by

$$[\mathcal{B}^\pi_{\mathbb{C}}(f)] = \begin{bmatrix} [\mathcal{B}'_1\pi] \\ \vdots \\ [\mathcal{B}'^\pi_\sigma] \end{bmatrix}, \text{ where } [\mathcal{B}'^\pi_i] = \begin{bmatrix} \gamma^i_1 & \delta^i_1 \\ \gamma^i_2 & \delta^i_2 \\ \vdots & \vdots \\ \gamma^i_{\nu_i} & \delta^i_{\nu_i} \end{bmatrix} \in \mathbb{C}^{\nu_i\times 2} \tag{6.18a}$$

The matrix $[\mathcal{B}^\pi_{\mathbb{C}}(f)]$ will be referred to as a ($\mathbb{C}$-π)-basis matrix of $f(s,\hat{s})$.

(iv) For the set $\mathcal{D}_{\mathbb{R}}(f)$ we may define in a similar manner the sets $\mathcal{B}_j \overset{\Delta}{=} \{(\alpha^i_j,\beta^i_j;\tau_i):\alpha^i_j,\beta^i_j\in\mathbb{R},\tau_i\in\mathbb{Z},j\in\underline{\mu}_i,(\alpha^i_j,\beta^i_j)\neq\zeta(\alpha^i_k,\beta^i_k), \forall\, j\neq k, \zeta\in\mathbb{R}-\{0\}\}$, $\mathcal{B}_{\mathbb{R}}(f)=\{\mathcal{B}_1,\dots,\mathcal{B}_\rho;\tau_1<\tau_2<\dots<\tau_\rho\}$ and $J_{\mathbb{R}}(f)=\{(\tau_1,\mu_1),\dots,(\tau_\rho,\mu_\rho)\}$, as well as the notions of normal ordering and of the matrix representation. The sets $\mathcal{B}_{\mathbb{R}}(f),J_{\mathbb{R}}(f)$ will be referred to as the real unique factorisation set ($\mathbb{R}$-UFS), real list respectively of $f(s,\hat{s})$ and the matrix $[\mathcal{B}^\pi_{\mathbb{R}}(f)]$ defined as in (6.18a) for some permutation π will be called an ($\mathbb{R}$-π)-basis matrix of $f(s,\hat{s})$.

(v) The sets $\mathcal{B}(f)=\{\mathcal{B}_{\mathbb{R}}(f);\mathcal{B}_{\mathbb{C}}(f)\}$ and $J(f)=\{J_{\mathbb{R}}(f);J_{\mathbb{C}}(f)\}$ will be called the unique factorisation set (UFS) and the list of $f(s,\hat{s})$ correspondingly. For every $\pi\in<\mathcal{B}_{\mathbb{R}}(f)>$ and $\pi'\in<\mathcal{B}_{\mathbb{C}}(f)>$ a matrix representation of $\mathcal{B}(f)$ according to (π,π') is defined by

$$[\mathcal{B}^{\pi,\pi'}(f)] = \begin{bmatrix} [\mathcal{B}^{\pi}_{\mathbb{R}}(f)] \\ [\mathcal{B}^{\pi'}_{\mathbb{C}}(f)] \end{bmatrix} \tag{6.18b}$$

$[\mathcal{B}^{\pi,\pi'}(f)]$ will be referred to as a <u>(π,π')-basis matrix</u> of $f(s,\hat{s})$.

<u>Remark (6.2)</u>: From the above definitions it is clear that since $\#\mathcal{B}_i^{!}=\nu_i$, $\#\mathcal{B}_i^{'}=\mu_i$, then $\#<\mathcal{B}_i^{!}>=\nu_i!$, $\#<\mathcal{B}_i>=\mu_i!$ and thus $\#<\mathcal{B}_{\mathbb{C}}(f)>=\nu_1!\nu_2!\ldots\nu_\sigma!$, $\#<\mathcal{B}_{\mathbb{R}}(f)>=\mu_1!\mu_2!\ldots\mu_\rho!$ and $\#<\mathcal{B}(f)>=\nu_1!\ldots\nu_\sigma!\mu_1!\ldots\mu_\rho!$.

With this notation in mind we may state the main result of this section.

<u>Theorem (6.2)</u>: Let $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})\in\mathbb{R}_d\{\Theta\}$ and let $\{\mathcal{B}_{\mathbb{R}}(f),J_{\mathbb{R}}(f),\mathcal{B}_{\mathbb{C}}(f),J_{\mathbb{C}}(f)\}$, $\{\mathcal{B}_{\mathbb{R}}(\tilde{f}),J_{\mathbb{R}}(\tilde{f}),\mathcal{B}_{\mathbb{C}}(\tilde{f}),J_{\mathbb{C}}(\tilde{f})\}$ be the corresponding sets associated with $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})$. $f(s.\hat{s})E_{\mathcal{P}}\tilde{f}(\lambda,\hat{\lambda})$, if and only if the following conditions hold true:

(i) $J_{\mathbb{R}}(f)=J_{\mathbb{R}}(\tilde{f})$ and $J_{\mathbb{C}}(f)=J_{\mathbb{C}}(\tilde{f})$.

(ii) There exist $\pi(\mathcal{B}_{\mathbb{R}}(f))\in<\mathcal{B}_{\mathbb{R}}(f)>,\tilde{\pi}(\mathcal{B}_{\mathbb{R}}(\tilde{f}))\in<\mathcal{B}_{\mathbb{R}}(\tilde{f})>,\pi'(\mathcal{B}_{\mathbb{C}}(f))\in<\mathcal{B}_{\mathbb{C}}(f)>$, $\tilde{\pi}'(\mathcal{B}_{\mathbb{C}}(\tilde{f}))\in<\mathcal{B}_{\mathbb{C}}(\tilde{f})>$, a $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$, $\zeta_i\in\mathbb{R}-\{0\}$ and $\xi_i\in\mathbb{C}-\{0\}$ such that

$$[\mathcal{B}^{\tilde{\pi}}_{\mathbb{R}}(\tilde{f})] = \mathrm{diag}\{\zeta_i\}[\mathcal{B}^{\pi}_{\mathbb{R}}(f)][\beta] \tag{6.19a}$$

$$[\mathcal{B}^{\tilde{\pi}'}_{\mathbb{C}}(\tilde{f})] = \mathrm{diag}\{\xi_i\}[\mathcal{B}^{\pi'}_{\mathbb{C}}(f)][\beta] \tag{6.19b}$$

□

The proof of the above result follows immediately by Propositions (6.5) and (6.6) and the fact that the pair (γ,δ) characterising an e.d. is defined modulo some nonzero constant.

<u>Corollary (6.1)</u>: The real and complex lists $J_{\mathbb{R}}(f),J_{\mathbb{C}}(f)$ of $f(s,\hat{s})$ are invariants of the $E_{\mathcal{P}}(f)$ equivalence class. □

Two pairs of sets $\mathcal{B}(f)=\{\mathcal{B}_{\mathbb{R}}(f),\mathcal{B}_{\mathbb{C}}(f)\}$, $\mathcal{B}(\tilde{f})=\{\mathcal{B}_{\mathbb{R}}(\tilde{f}),\mathcal{B}_{\mathbb{C}}(\tilde{f})\}$ for which $J_{\mathbb{R}}(f)=J_{\mathbb{R}}(\tilde{f}),J_{\mathbb{C}}(f)=J_{\mathbb{C}}(\tilde{f})$ and conditions (ii) of Theorem (6.2) hold true for some $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ and nonzero constants ζ_i,ξ_i, will be called <u>normally projective equivalent</u> (NPE) and shall be denoted by $\mathcal{B}(f)E_{\mathcal{P}}\mathcal{B}(\tilde{f})$.

Clearly, this notion of equivalence defined now on the sets $\mathcal{B}(f)$ is equivalent to the $\mathcal{E}_{\mathcal{P}}$-equivalence notion defined on $\mathbb{R}_d\{\Theta\}$. In other words, $\mathcal{E}_{\bar{\mathcal{P}}}$-equivalence is nothing else but $\mathcal{E}_{\mathcal{P}}$-equivalence defined on the UFS of the polynomials in $\mathbb{R}_d\{\Theta\}$. The advantage of the $\mathcal{E}_{\bar{\mathcal{P}}}$-equivalence notion is that the study of invariants of $\mathcal{E}_{\mathcal{P}}(f)$ is reduced to a standard matrix algebra problem, i.e. the study of solvability of condition (6.19a,b).

Before we proceed to the investigation of the conditions for the solvability of (6.19a),(6.19b), it is worth pointing out that the study of $\mathcal{E}_{\mathcal{P}}$-equivalence, as it has been expressed by Theorem (6.2), is equivalent to the following problem of algebraic projective geometry: Given two symmetric sets of points $\zeta,\tilde{\zeta}$ of $\mathbb{C}\cup\{\infty\}$, find the necessary and sufficient conditions for the existence of a projective transformation $\beta\epsilon PGL(1,\mathbb{C}/\mathbb{R})$ such that ζ is mapped to $\tilde{\zeta}$ under β; this problem will be referred to as the general linear mapping problem (GLMP) on $\mathbb{C}\cup\{\infty\}$. In our attempt to characterise $\mathcal{E}_{\mathcal{P}}(f)$ by a complete set of invariants, it is clear that we have to solve the GLMP. Apart from some general necessary conditions [Se & Kn - 1], characterising the solvability of GLMP, the set of necessary and sufficient conditions characterising the solvability of GLMP have not been worked out before. Central to our attempt to solve the GLMP and thus find a complete set of invariants for $\mathcal{E}_{\mathcal{P}}(f)$ is the study of the notion of extended Hermite equivalence of matrices discussed next.

6.5 Extended Hermite equivalence of matrices

6.5.1 Statement of the problem and background definitions

The central problem in the study of invariants of $\mathcal{E}_{\mathcal{P}}(f)$, or the study of GLMP is the investigation of the solvability of conditions (6.19a,b); these conditions express a notion of equivalence defined on matrices and it is studied here.

Let $T\epsilon\mathbb{C}^{k\times 2}$. The matrix T will be called entirely nonsingular if none of its 2×2 minors is zero. The set of entirely nonsingular matrices of k×2 dimensions shall be denoted by $\mathbb{C}_n^{k\times 2}$; the subset of $\mathbb{C}_n^{k\times 2}$, defined by those matrices T having real elements, will be denoted by $\mathbb{R}_n^{k\times 2}$. Clearly, $[\mathcal{B}_{\mathbb{C}}^{\pi'}(f)]$, $[\mathcal{B}_{\mathbb{R}}^{\pi}(f)]$ are entirely nonsingular matrices (complex, real respectively).

Definition (6.2): Let $T_1, T_2 \in \mathbb{C}^{k\times 2}$. T_1, T_2 will be called complex extended Hermite equivalent and shall be denoted by $T_1 \mathcal{E}_{eh}^{c} T_2$, iff there exist $\xi_i \in \mathbb{C}-\{0\}, i\in\underset{\sim}{k}$ and $Q\in\mathbb{C}^{2\times 2}$, $|Q|=\delta\in\mathbb{R}-\{0\}$ such that

$$T_2 = \operatorname{diag}\{\xi_i\} T_1 Q \tag{6.20}$$

T_1, T_2 will be called real extended Hermite equivalent, and shall be denoted by $T_1 \mathcal{E}_{eh}^{r} T_2$, iff (6.20) holds true for some $Q\in\mathbb{R}^{2\times 2}$, $|Q|=\delta\in\mathbb{R}-\{0\}$.

Remark (6.3): In the case of $\mathcal{E}_{eh}^{c}$-equivalence we may always assume that $|Q|=\delta=1$, whereas in the case of $\mathcal{E}_{eh}^{r}$-equivalence, we may always assume that $|Q|=\delta=1$, or -1. This is due to the fact that if $\delta\neq(\pm 1)$, then $|\delta|$ may be incorporated into the parameters ξ_i.

Clearly, the study of $\mathcal{E}_{\bar{p}}$-equivalence of sets $\mathcal{B}(f)$ is reduced to an equivalent study of $\mathcal{E}_{eh}^{r}$-equivalence defined on the corresponding $[\mathcal{B}^{\pi,\pi'}(f)]$ entirely nonsingular matrix. Given that $\mathcal{E}_{eh}^{r}$-equivalence is a special case of $\mathcal{E}_{eh}^{c}$-equivalence, the more general notion of $\mathcal{E}_{eh}^{c}$-equivalence is considered first on the set $\mathbb{C}_n^{k\times 2}$, and then the results will be specialised to the $\mathcal{E}_{eh}^{r}$-equivalence case.

The main problem considered here is the definition of a complete set of invariants of $T\in\mathbb{C}_n^{k\times 2}$ matrices under $\mathcal{E}_{eh}^{c}$-equivalence first and then under $\mathcal{E}_{eh}^{r}$-equivalence. Such a study involves the use of some extra notation and definitions from exterior algebra.

Definition (6.3): Let $A=[\underline{a}_1,\dots,\underline{a}_m]\in\mathbb{C}^{n\times m}$, $m+2\leq n$ and let $\underline{g}(A)=\underline{a}_1\wedge\dots\wedge\underline{a}_m=$ $=(a_\omega)\in\mathbb{C}^{\binom{n}{m}}$, $\omega\in Q_{m,n}$ be the exterior product of the columns of A (Grassmann vector), where a_{ω} denote the coordinates of $\underline{g}(A)$.

(i) [Mar - 1] A *triangle* of $Q_{m,n}$, based on $\phi=\{i_1,i_1,i_m,i_{m+1},i_{m+2}\}\in Q_{m+2,n}$, i.e. i_j $\{1,2,\dots,n\}$, $i_1<i_2<\dots<i_m<i_{m+1}<i_{m+2}$, is a set of six sequences in $Q_{m,n}$ of the following form:

$$\delta\phi=\{(i_1,i_2,\dots,i_{m-1},i_m),(i_1,i_2,\dots,i_{m-1},i_{m+1}),(i_1,i_2,\dots,i_{m-1},i_{m+2}),$$

$$(i_2,i_3,\dots,i_{m-1},i_{m+1},i_{m+2}),(i_2,i_3,\dots,i_{m-1},i_m,i_{m+2}),(i_2,i_3,\dots,i_{m-1},i_m,i_{m+1})\}$$

(ii) A *Plücker vector* of A based on the triangle defined by ϕ is defined by

$$\underline{p}_A(\phi)=\begin{bmatrix} a(i_1,i_2,\dots,i_{m-1},i_m)\ \ a(i_2,i_3,\dots,i_{m-1},i_{m+1},i_{m+2} \\ -a(i_1,i_2,\dots,i_{m-1},i_{m+1})a(i_2,i_3,\dots,i_{m-1},i_m,i_{m+2}) \\ a(i_1,i_2,\dots,i_{m-1},i_{m+2})a(i_2,i_3,\dots,i_{m-1},i_m,i_{m+1}) \end{bmatrix}\in\mathbb{C}^3$$

(iii) Let $<k>=\{1,2,\dots,k\}$ be the k-set of positive integers and let $r\in<k>$ be a fixed integer. The set of quadraples

$$\Phi_r \overset{\Delta}{=} \{\phi_r:\phi_r=(i_1,i_2,i_3,i_4)\in Q_{4,k} \text{ for which } r\in\{i_1,i_2,i_3,i_4\}\}$$

will be called the set of prime quadraple of $r\in<k>$. For each prime quadraple $\phi_r\in\Phi_r$ we can consider the corresponding triangle, $\delta\phi_r$, of sequences of $Q_{2,k}$ based on ϕ_r; the set of all triangles based on all prime quadraples of r will be called the set of *prime triangles* of r and shall be denoted by $\Delta\Phi_r$. Thus,

$$\Delta\phi_r=\{\delta\phi_r:\delta\phi_r=\{(i_1,i_2),(i_1,i_3),(i_1,i_4),(i_2,i_3),(i_2,i_4),(i_3,i_4)\}$$
$$\text{for } \forall\ \phi_r=(i_1,i_2,i_3,i_4)\in\Phi_r\}$$

(iv) If $T=[\underline{t}_1,\underline{t}_2]\in\mathbb{C}^{k\times 2}$, $k\geq 4$, $\underline{g}(T)=\underline{t}_1\wedge\underline{t}_2=(a_{\omega})\in\mathbb{C}^{\binom{k}{2}}$, $\omega=(i_1,i_2)\in Q_{2,k}$, and $r\in<k>$ fixed, then we define as the set of *r-prime Plücker vectors* of T, the set of vectors of $\mathbb{C}^3$ defined by:

$$\mathcal{P}_r \triangleq \{ \underline{p}(\delta\phi_r) : \underline{p}(\delta\phi_r) = \begin{bmatrix} a(i_1,i_2)a(i_3,i_4) \\ -a(i_1,i_3)a(i_2,i_4) \\ a(i_1,i_4)a(i_2,i_3) \end{bmatrix} , \forall\, \delta\phi_r \in \Delta\Phi_r \}$$

□

We may illustrate the above definition by the following examples:

Example (6.1): Let $T \in \mathbb{C}^{5\times 2}$, and let $r=1 \in <5> = \{1,2,3,4,5\}$.

(i) The set of prime quadraples of $r=1$ is given by:

$$\Phi_1 = \{(1,2,3,4),(1,2,3,5),(1,2,4,5),(1,3,4,5)\} = \{\phi_1^1,\phi_1^2,\phi_1^3,\phi_1^4\}$$

(ii) The set of prime triangles of $r=1$ is given by $\Delta\Phi_1 = \{\delta\phi_1^1,\delta\phi_1^2,\delta\phi_1^3,\delta\phi_1^4\}$ where

$$\delta\phi_1^1 = \{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\},$$

$$\delta\phi_1^2 = \{(1,2),(1,3),(1,5),(2,3),(2,5),(3,5)\},$$

$$\delta\phi_1^3 = \{(1,2),(1,4),(1,5),(2,4),(2,5),(4,5)\},$$

$$\delta\phi_1^4 = \{(1,3),(1,4),(1,5),(3,4),(3,5),(4,5)\}.$$

(iii) If $T \in C^{5\times 2}$ and $\underline{g}(T) = (a_\omega) \in \mathbb{C}^{\binom{5}{2}}$, $\omega = (i_1,i_2) \in Q_{2,5}$, where a_ω is the 2×2 minor of T that corresponds to the (i_1,i_2) rows of T, then the set of 1-prime vectors is defined by

$$\mathcal{P}_1 = \{ \begin{bmatrix} a(1,2)a(3,4) \\ -a(1,3)a(2,4) \\ a(1,4)a(2,3) \end{bmatrix}, \begin{bmatrix} a(1,2)a(3,5) \\ -a(1,3)a(2,5) \\ a(1,5)a(2,3) \end{bmatrix}, \begin{bmatrix} a(1,2)a(4,5) \\ -a(1,4)a(2,5) \\ a(1,5)a(2,4) \end{bmatrix}, \begin{bmatrix} a(1,3)a(4,5) \\ -a(1,4)a(3,5) \\ a(1,5)a(3,4) \end{bmatrix} \}$$

□

Remark (6.4): Note that a triangle of sequences of $Q_{m,n}$ may be used for the derivation of the three term Quadratic Plücker relationships (QPR), characterising decomposability of multivectors [Mar-1]. By construction $\underline{g}(T)$ is decomposable, and thus the sum of the coordinates for every Plücker vector is equal to zero (a_ω satisfy the three term QPR), i.e.

$a(1,2)a(3,4)-a(1,3)a(2,4)+a(1,4)a(2,3)=0$. The term "Plücker vectors" has been used because of the links of these vectors with the set of 3-term QPR's.

6.5.2 The Plücker vectors as a complete set of invariants for matrices of $\mathbb{C}_n^{k\times 2}$ under E^c_{eh}-equivalence

With those preliminary definitions in mind we may proceed to the investigation of the conditions under which two matrices of $\mathbb{C}_n^{k\times 2}$ are E^c_{eh}-equivalent. This study eventually yields a complete set of invariants for the equivalence class $E^c_{eh}(T)$, where $T\in\mathbb{C}_n^{k\times 2}$. We start off with the following obvious result:

Proposition (6.7): Let $T_1,T_2\in\mathbb{C}^{k\times 2}$. Necessary and sufficient conditions for $T_1 E^r_{eh} T_2$ are:

(i) There exist $\xi_i\in\mathbb{C}-\{0\}$ such that

$$\text{col-span}_{\mathbb{C}}\{T_2\} = \text{col-span}_{\mathbb{C}}\{\text{diag}\{\xi_i\}T_1\} \tag{6.21}$$

(ii) For at least a set of solutions $\{\xi_i\}$ of (6.21), the following two conditions hold true:

$$\text{col-span}_{\mathbb{R}}\{\text{Re}(T_2)\} = \text{col-span}_{\mathbb{R}}\{\text{Re}(\text{diag}\{\xi_i\}T_1)\} \tag{6.22a}$$

$$\text{col-span}_{\mathbb{R}}\{\text{Im}(T_2)\} = \text{col-span}_{\mathbb{R}}\{\text{Im}(\text{diag}\{\xi_i\}T_1)\} \tag{6.22b}$$

□

Part (i) of the above result expresses the necessary and sufficient conditions for $T_1 E^c_{eh} T_2$. Part (ii) is the extra condition needed to guarantee the existence of a real Q. In some sense, Proposition (6.7) is a restatement of the notion of E^r_{eh}-equivalence. The fact that conditions (6.21) and (6.22) depend on the set of $\{\xi_i\}$ does not make this result particularly useful; however, Proposition (6.7) is the starting point in the search for a complete set of invariants.

Remark (6.5): Condition (6.21) is necessary and sufficient for $\mathcal{E}^c_{eh}$-equivalence of matrices of $\mathbb{C}^{k\times 2}$. If $T_1,T_2\in\mathbb{R}^{k\times 2}$, then (6.21) is a necessary and sufficient condition for $\mathcal{E}^r_{eh}$-equivalence.

In the following, attention is focussed on $\mathcal{E}^c_{eh}$-equivalence of matrices of $\mathbb{C}^{k\times 2}_n$. By Proposition (6.7) we have:

Proposition (6.8): Let $T_1=[\underline{t}_{11},\underline{t}_{12}], T_2=[\underline{t}_{21},\underline{t}_{22}]\in\mathbb{C}^{k\times 2}_n, k\geq 2$, and let $\underline{t}_{11}\wedge\underline{t}_{12}=(a^1_\omega), \underline{t}_{21}\wedge\underline{t}_{22}=(a^2_\omega)\in\mathbb{C}^{\binom{k}{2}}, \omega\in Q_{2,k}$, where a^1_ω, a^2_ω are the 2×2 minors of T_1,T_2 respectively, which correspond to the $\omega=(i_1,i_2)$ rows, and let

$$a^2_\omega/a^1_\omega=\xi_\omega\delta,\ \xi_\omega\in\mathbb{C}-\{0\},\delta\in\mathbb{R}-\{0\} \tag{6.23a}$$

Necessary and sufficient condition for $T_1\mathcal{E}^c_{eh}T_2$ is that there exists a set of $\xi_i\in\mathbb{C}-\{0\}, i\in\underset{\sim}{k}$, such that

$$\xi_\omega\ =\xi_{i_1}\xi_{i_2}\ \text{ for }\forall\ \omega=(i_1,i_2)\in Q_{2,k} \tag{6.23b}$$

Proof

By taking the second compound matrix of both sides of (6.20) and by using the Binet-Cauchy Theorem [Mar & Min - 1], we have

$$\underline{\underline{g}}(T_2)=C_2(T_2)=C_2(\text{diag}\{\xi_i\}T_1Q)=C_2(\text{diag}\{\xi_i\})\underline{\underline{g}}(T_1)C_2(Q) \tag{6.24a}$$

Note that $C_2(Q)=|Q|=\delta\neq 0$ and that $C_2(\text{diag}\{\xi_i\})=\text{diag}\{\xi_\omega\}$, where $\xi_\omega=\xi_{i_1}\xi_{i_2}$, $\omega=(i_1,i_2)\in Q_{2,k}$. Thus, by (6.24a) we have

$$\begin{bmatrix}\vdots\\ a^2_\omega\\ \vdots\end{bmatrix}=\text{diag}\{\xi_\omega\}\begin{bmatrix}\vdots\\ a^1_\omega\\ \vdots\end{bmatrix}\delta,\ \xi_\omega=\xi_{i_1}\xi_{i_2},\delta\in\mathbb{R}-\{0\} \tag{6.24b}$$

and the necessity is established. To prove sufficiency, we note that if $\xi_\omega=\xi_{i_1}\xi_{i_2}$ for $\forall\,\omega=(i_1,i_2)\in Q_{2,k}$ and some k parameters ξ_i, then $\text{diag}\{\xi_{i_1}\xi_{i_2}\}=\text{diag}\{\xi_\omega\}=C_2(\text{diag}\{\xi_i\})$ and because the multivectors $(a^2_\omega),(a^1_\omega)$ are decomposable we may write (6.24b) as

$$\underline{g}(T_2) = C_2(\text{diag}\{\xi_i\})\underline{g}(T_1)\delta = \underline{g}(\text{diag}\{\xi_i\}T_1)\delta \qquad (6.24c)$$

From the above it follows that T_2 and $\text{diag}\{\xi_i\}T_1$ have the same column space (they are Grassmann representatives of the same space) and thus there exists a $Q\in\mathbb{C}^{2\times 2}$, $|Q|=\delta\neq 0$ for which $T_2=\text{diag}\{\xi_i\}T_1Q$. □

Remark (6.6): Products $\xi_\omega\delta$ are uniquely defined by eqn(6.23a). For the case of $\mathcal{E}^c_{eh}$-equivalence, if $\delta\neq 1$ then

$$T_2 = \text{diag}\{\xi_i\}T_1Q \iff T_2 = \text{diag}\{\sqrt{\delta}\xi_i\}T_1\{Q.1/\sqrt{\delta}\} \qquad (6.25a)$$

where $|Q.1/\sqrt{\delta}|=1$. Thus for $\mathcal{E}^c_{eh}$, δ may be assumed to be 1. For the case of $\mathcal{E}^r_{eh}$-equivalence, if $\delta\neq 1$, then

$$T_2 = \text{diag}\{\xi_i\}T_1Q \iff T_2 = \text{diag}\{\sqrt{|\delta|}\xi_i\}T_1\{Q.1\sqrt{|\delta|}\} \qquad (6.25b)$$

where $|Q.1\sqrt{|\delta|}|=1$ if $\delta>0$ and $|Q.1\sqrt{|\delta|}|=-1$ if $\delta<0$. Thus, for $\mathcal{E}^r_{eh}$-equivalence, δ may be assumed to be either 1, or -1. □

Note that Proposition (6.8) provides the necessary and sufficient conditions for the existence of complex scalars ξ_i and a real determinant complex transformation Q for which $T_2=\text{diag}\{\xi_i\}T_1Q$; however, this result does not guarantee the existence of a real Q. In the case, however, where T_1 and T_2 are real, then Proposition (6.8) guarantees the existence of a real Q.

The essence of Proposition (6.8) is that it reduces the study of $\mathcal{E}^c_{eh}$-equivalence to a factorisation problem of the set of ξ_ω defined by (6.23a) into the form $\xi_\omega=\xi_{i_1}\xi_{i_2}$ for all $\omega\in Q_{2,k}$, in terms of the k-parameters ξ_i. Such a factorisation is necessary and sufficient for the matrix $\text{diag}\{\xi_\omega\}=\text{diag}\{a^2_\omega/a^1_\omega.1/\delta\}\in\mathbb{C}^{\binom{k}{2}\times\binom{k}{2}}$ to be considered as the second order compound matrix of some $\text{diag}\{\xi_i\}$ matrix. Thus the study of $\mathcal{E}^c_{eh}$-equivalence may be expressed as a problem of decomposability of linear operators. We may summarise the discussion as follows:

Proposition (6.9):

(i) Necessary condition for $T_1 \mathcal{E}^r_{eh} T_2$ is that the set of equations

$$\delta\xi_{i_1}\xi_{i_2} = a^{(2)}_\omega / a^{(1)}_\omega, \ \omega = (i_1, i_2) \in Q_{2,k} \tag{6.26a}$$

have a solution $\{\xi_1, \xi_2, \ldots, \xi_k : \xi_i \in \mathbb{C}-\{0\}\}$ for $\delta=1$ or $\delta=-1$.

(ii) Necessary and sufficient condition for $T_1 \mathcal{E}^c_{eh} T_2$ is that the set of equations

$$\xi_{i_1}\xi_{i_2} = a^{(2)}_\omega / a^{(1)}_\omega, \ \omega = (i_1, i_2) \in Q_{2,k} \tag{6.26b}$$

have a solution $\{\xi_1, \xi_2, \ldots, \xi_k : \xi_i \in \mathbb{C}-\{0\}\}$. □

Before we examine the conditions under which a solution of (6.26a) and (6.26b) exists we give a result expressing the uniqueness property of such solutions.

Proposition (6.10):

(i) If $\{\xi_i : \xi_i \in \mathbb{C}-\{0\}, i \in \underset{\sim}{k}\}$ and $\{\xi'_i : \xi'_i \in \mathbb{C}-\{0\}, i \in \underset{\sim}{k}\}$ are two solutions of (6.26a) for $\delta=1$ or $\delta=-1$, then either $\xi'_i = \xi_i$ for $\forall\ i \in \underset{\sim}{k}$ or $\xi'_i = j\xi_i$ for $\forall\ i \in \underset{\sim}{k}$.

(ii) If $\{\xi_i : \xi_i \in \mathbb{C}-\{0\}, i \in \underset{\sim}{k}\}$ is a solution of (6.26b), then it is uniquely defined.

Proof

Since $\{\xi_i\}, \{\xi'_i\}$ are two solutions, then for any three indices i_1, i_2, i_3 we have that

$$a^2_{i_1,i_2} / a^1_{i_1,i_2} = \xi_{i_1}\xi_{i_2}\delta = \xi'_{i_1}\xi'_{i_2}\delta' \Rightarrow \xi_{i_1}/\xi'_{i_1} = (\xi'_{i_2}/\xi_{i_2}).(\delta'/\delta)$$

$$a^2_{i_1,i_3} / a^1_{i_1,i_3} = \xi_{i_1}\xi_{i_2}\delta = \xi'_{i_1}\xi'_{i_2}\delta' \Rightarrow \xi_{i_1}/\xi'_{i_1} = (\xi'_{i_3}/\xi_{i_3}).(\delta'/\delta)$$

$$a^2_{i_2,i_3} / a^1_{i_2,i_3} = \xi_{i_2}\xi_{i_3}\delta = \xi'_{i_2}\xi'_{i_3}\delta' \Rightarrow \xi_{i_2}/\xi'_{i_2} = (\xi'_{i_3}/\xi_{i_3}).(\delta'/\delta)$$

from which it is readily seen that

$$\xi_{i_1}/\xi'_{i_1} = (\xi'_{i_2}/\xi_{i_2})(\delta'/\delta) = (\xi'_{i_3}/\xi_{i_3})(\delta'/\delta) = \xi_{i_2}/\xi'_{i_2}$$

or that $(\xi'_{i_2}/\xi_{i_2})^2=\delta/\delta'$. Similarly, it can be proved that

$$\xi'_j/\xi_j = \sqrt{\delta/\delta'} \text{ for all } j\in\underset{\sim}{k}$$

and since $\delta=\pm 1$ the result follows. In the case eqn(6.26b), $\delta=\delta'=1$ and from the last expression the result is established. □

Remark (6.7) clearly demonstrates that the solutions of (6.26a),(6.26b) do not possess the properties described by Proposition (6.10), if $\delta\neq 1$. In the following, it will be assumed that $\delta=1$, in the case of $\mathcal{E}^c_{eh}$-equivalence, and $\delta=1$ or -1, in the case of $\mathcal{E}^r_{eh}$-equivalence. The conditions under which $T_1\mathcal{E}^c_{eh}T_2$ are examined next; we start off by considering the simple cases where k=1,2,3.

Proposition (6.11): Let $T_1,T_2\in\mathbb{C}^{k\times 2}_n$, where k=1,2,3. Then $T_1\mathcal{E}^c_{eh}T_2$.

Proof

The proof for k=1,2 is rather obvious and it is omitted. For the case k=3, conditions (6.26a) may be solved. Thus we have

$$\xi_1 = \frac{1}{\sqrt{\delta}}\frac{a^1_{2,3}}{a^2_{2,3}}\sqrt{D},\ \xi_2 = \frac{1}{\sqrt{\delta}}\frac{a^1_{1,3}}{a^2_{2,3}}\sqrt{D},\ \xi_3 = \frac{1}{\sqrt{\delta}}\frac{a^1_{1,2}}{a^2_{1,2}}\sqrt{D} \tag{6.27a}$$

where $\delta=1$, or -1, and

$$D = (a^2_{1,2}a^2_{1,3}a^2_{2,3})/(a^1_{1,2}a^1_{1,3}a^1_{2,3}) \tag{6.27b}$$

Therefore, for all $T_1,T_2\in\mathbb{C}^{3\times 2}_n$, we can always find two solutions (ξ_1,ξ_2,ξ_3) and $(j\xi_1,j\xi_2,j\xi_3)$ where the first corresponds to a $Q\in\mathbb{C}^{2\times 2}$ with $|Q|=1$ and the second to a Q with $|Q|=-1$. □

Remark (6.7): Any set of three distinct real points may be mapped by a projective transformation $\beta\in PGL(1,\mathbb{C}|\mathbb{R})$ to any three real distinct points.

We may now state the main result of this section.

<u>Theorem (6.3)</u>: Let $T_1,T_2\in\mathbb{C}_n^{k\times 2},k\geq 4,\underline{g}(T_1)=(a_\omega^1),\underline{g}(T_2)=(a_\omega^2)\in\mathbb{C}_n^{\binom{k}{2}}$ be the corresponding Grassmann vectors, and $a_\omega^1,a_\omega^2,\omega\in Q_{2,k}$ be the respective Plücker coordinates. Let us further assume that $r\in\langle k\rangle$, is a fixed integer, Φ_r and $\Delta\Phi_r$ are the sets of prime quadruples and prime triangles of r and that $\mathcal{P}_r^1,\mathcal{P}_r^2$ are the r-prime Plücker vectors of T_1,T_2 respectively. Necessary and sufficient condition for $T_1\mathcal{E}_{eh}^c T_2$ are that for $\forall\ \delta\phi_r\in\Delta\Phi_r$ and $\underline{p}^1(\delta\phi_r)\in\mathcal{P}_r^1$, $\underline{p}^2(\delta\phi_r)\in\mathcal{P}_r^2$

$$\underline{p}^2(\delta\phi_r)=\lambda_j\underline{p}^1(\delta\phi_r)\text{ for some }\lambda_j\in\mathbb{C}-\{0\}\tag{6.28}$$

<u>Proof</u>

We shall prove the result for r=1, whereas the proof for any general $r\in\langle k\rangle$ is identical. We consider first the cases where k=4,k=5 and by induction we shall prove the general case. In the following, we shall denote by $q_\omega=a_\omega^2/a_\omega^1,\hat{q}_\omega=a_\omega^1/a_\omega^2$, for $\forall\ \omega=(i_1,i_2)\in Q_{2,k}$.

(i) <u>k=4</u>: Conditions (6.26a) yield

$$\delta\xi_1\xi_2=q_{1,2},\ \delta\xi_1\xi_3=q_{1,3},\ \delta\xi_1\xi_4=q_{1,4}\tag{6.29a}$$

$$\delta\xi_2\xi_3=q_{2,3},\ \delta\xi_2\xi_4=q_{2,4},\ \delta\xi_3\xi_4=q_{3,4}\tag{6.29b}$$

Solving (6.29a) for ξ_2,ξ_3,ξ_4 in terms of $\delta\in\mathbb{R}-\{0\}$ and $\xi_1\in\mathbb{C}-\{0\}$ and substituting into (6.29b), we have the equivalent set

$$\xi_2=q_{1,2}/\delta\xi_1,\ \xi_3=q_{1,3}/\delta\xi_1,\ \xi_4=q_{1,4}/\delta\xi_1\tag{6.30a}$$

$$q_{1,2}q_{1,3}\hat{q}_{2,3}=q_{1,2}q_{1,4}\hat{q}_{2,4}=q_{1,3}q_{1,4}\hat{q}_{3,4}=\delta\xi_1^2\tag{6.30b}$$

However, (6.30b) is equivalent to the following conditions

$$q_{1,2}q_{1,3}\hat{q}_{2,3}=q_{1,2}q_{1,4}\hat{q}_{2,4}\iff q_{1,3}q_{2,4}=q_{1,4}q_{2,3}=\lambda\tag{6.30c}$$

$$q_{1,2}q_{1,3}\hat{q}_{2,3}=q_{1,3}q_{1,4}\hat{q}_{3,4}\iff q_{1,2}q_{3,4}=q_{1,4}q_{2,3}=\lambda\tag{6.30d}$$

from which

$$\begin{bmatrix} a^2_{1,2} & a^2_{3,4} \\ -a^2_{1,3} & a^2_{2,4} \\ a^2_{1,4} & a^2_{2,3} \end{bmatrix} = \lambda \begin{bmatrix} a^1_{1,2} & a^1_{3,4} \\ -a^1_{1,3} & a^1_{2,4} \\ a^1_{1,4} & a^1_{2,3} \end{bmatrix} \text{ or } \underline{p}^2(1,2,3,4) = \lambda\underline{p}^1(1,2,3,4) \qquad (6.30e)$$

and thus necessity is proved. The sufficiency follows by a mere reversion of the steps.

(ii) <u>k=5</u>: There are ten conditions of the (6.26a) type. By solving the first four for ξ_2,ξ_3,ξ_4,ξ_5 in terms of δ,ξ_1 and substituting into the rest, the following equivalent set is obtained

$$\xi_2 = q_{1,2}/\delta\xi_1,\ \xi_3 = q_{1,3}/\delta\xi_1,\ \xi_4 = q_{1,4}/\delta\xi_1,\ \xi_5 = q_{1,5}/\delta\xi_1 \qquad (6.31a)$$

$$q_{1,2}q_{1,3}\hat{q}_{2,3} = q_{1,2}q_{1,4}\hat{q}_{2,4} = q_{1,2}q_{1,5}\hat{q}_{2,5} = \delta\xi_1^2 =$$

$$= q_{1,3}q_{1,5}\hat{q}_{3,5} = q_{1,4}q_{1,5}\hat{q}_{4,5} \qquad (6.31b)$$

For the set <5>={1,2,3,4,5} the prime quadruples of 1 are {(1,2,3,4), (1,2,3,5),(1,2,4,5),(1,3,4,5)}; for each of those prime quadruples a three-term relationship is obtained from (6.31b), i.e.

$$(1,2,3,4) \rightarrow q_{1,2}q_{1,3}\hat{q}_{2,3} = q_{1,2}q_{1,4}\hat{q}_{2,4} = q_{1,3}q_{1,4}\hat{q}_{3,4} = \delta\xi_1^2 \qquad (6.31c)$$

$$(1,2,3,5) \rightarrow q_{1,2}q_{1,3}\hat{q}_{2,3} = q_{1,2}q_{1,5}\hat{q}_{2,5} = q_{1,3}q_{1,5}\hat{q}_{3,5} = \delta\xi_1^2 \qquad (6.31d)$$

$$(1,2,4,5) \rightarrow q_{1,2}q_{1,4}\hat{q}_{2,4} = q_{1,2}q_{1,5}\hat{q}_{2,5} = q_{1,4}q_{1,5}\hat{q}_{4,5} = \delta\xi_1^2 \qquad (6.31e)$$

$$(1,3,4,5) \rightarrow q_{1,3}q_{1,4}\hat{q}_{3,4} = q_{1,3}q_{1,5}\hat{q}_{3,5} = q_{1,4}q_{1,5}\hat{q}_{4,5} = \delta\xi_1^2 \qquad (6.31f)$$

Using similar arguments as in the case k=4, the above conditions yield

$$(6.31c) \iff \begin{bmatrix} a^2_{1,2}a^2_{3,4} \\ -a^2_{1,3}a^2_{2,4} \\ a^2_{1,4}a^2_{2,3} \end{bmatrix} = \lambda_1 \begin{bmatrix} a^1_{1,2}a^1_{3,4} \\ -a^1_{1,3}a^1_{2,4} \\ a^1_{1,4}a^1_{2,3} \end{bmatrix},\ (6.31d) \iff \begin{bmatrix} a^2_{1,2}a^2_{3,5} \\ -a^2_{1,3}a^2_{2,5} \\ a^2_{1,5}a^2_{2,3} \end{bmatrix} = \lambda_2 \begin{bmatrix} a^1_{1,2}a^1_{3,5} \\ -a^1_{1,3}a^1_{2,5} \\ a^1_{1,5}a^1_{2,3} \end{bmatrix}$$

$$(6.31e) \iff \begin{bmatrix} a^2_{1,2}a^2_{4,5} \\ -a^2_{1,4}a^2_{2,5} \\ a^2_{1,5}a^2_{2,4} \end{bmatrix} = \lambda_3 \begin{bmatrix} a^1_{1,2}a^1_{4,5} \\ -a^1_{1,4}a^1_{2,5} \\ a^1_{1,5}a^1_{2,4} \end{bmatrix}, (6.31f) \iff \begin{bmatrix} a^2_{1,3}a^2_{4,5} \\ -a^2_{1,4}a^2_{3,5} \\ a^2_{1,5}a^2_{3,4} \end{bmatrix} = \lambda_4 \begin{bmatrix} a^1_{1,3}a^1_{4,5} \\ -a^1_{1,4}a^1_{3,5} \\ a^1_{1,5}a^1_{3,4} \end{bmatrix} \tag{6.31g}$$

Note that (6.31g) is the minimum set of necessary conditions implied from eqn(6.31b). The sufficiency follows by a mere reversion of the arguments.

(iii) <u>The general case $k \geq 6$</u>: Before we proceed with the proof we define the following: Let $<k,1>=\{2,3,\dots,k\}$ and let $Q^{1)}_{2,k-1}$ be the set of strictly increasing 2 integers chosen from $<k,1>$ which are lexicographically ordered. We also define by $Q^{1]}_{3,k} \triangleq \{(1,i_1,i_2):(i_1,i_2)\in Q^{1)}_{2,k-1}\} \subset Q_{3,k}$; $Q^{1]}_{3,k}$ is the subset of $Q_{3,k}$ which contain 1 as an element (L is always first because of the lexicographic ordering of $Q_{3,k}$). By solving as before the first k-1 of conditions (6.26a) with respect to ξ_1, δ and substituting into the rest, we have the following set of equivalent conditions

$$\xi_2 = q_{1,2}/\delta\xi_1,\ \xi_3 = q_{1,3}/\delta\xi_1, \dots,\ \xi_k = q_{1,k}/\delta\xi_1 \tag{6.32a}$$

$$q_{1,i_1}q_{1,i_2}\hat{q}_{i_1,i_2} = \delta\xi_1^2 \text{ for all } (1,i_1,i_2)\in Q^{1]}_{3,k} \tag{6.32b}$$

If Φ_1 is the set of prime quadruples of (1), i.e. $\Phi_1=\{(1,i_2,i_3,i_4): i_2,i_3,i_4\in<k,1>, i_2<i_3<i_4\}$, then for each prime quadrupole $(1,i_2,i_3,i_4)$, a three-term relationship is obtained from (6.32b), of the type

$$q_{1,i_2}q_{1,i_3}\hat{q}_{i_2,i_3} = q_{1,i_2}q_{1,i_4}\hat{q}_{i_2,i_4} = q_{1,i_3}q_{1,i_4}\hat{q}_{i_3,i_4} = \delta\xi_1^2 \tag{6.32c}$$

The set of conditions (6.32c), defined for every $(1,i_2,i_3,i_4)\in\Phi_1$, is the minimal set of conditions needed for (6.32b) to be true. Following a similar procedure as in the (i),(ii) cases, it is readily shown that (6.32c) defined for a given $(1,i_2,i_3,i_4)\in\Phi_1$ is equivalent to

$$\begin{bmatrix} a^2_{1,i_2} & a^2_{i_3,i_4} \\ -a^2_{1,i_3} & a^2_{i_2,i_4} \\ a^2_{1,i_4} & a^2_{i_2,i_3} \end{bmatrix} = \lambda_{i_2 i_3 i_4} \begin{bmatrix} a^1_{1,i_2} & a^1_{i_3,i_4} \\ -a^1_{1,i_3} & a^1_{i_2,i_4} \\ a^1_{1,i_4} & a^1_{i_2,i_3} \end{bmatrix} \leftarrow (1,i_2,i_3,i_4)\in\Phi_1 \qquad (6.32d)$$

or that $\underline{p}^2(\delta\phi_1)=\lambda_j\underline{p}^1(\delta\phi_1)$ for $\forall\ \delta\phi_1\in\Delta\Phi_1$ and $\lambda_j\in\mathbb{C}-\{0\}$. The sufficiency follows by a mere reversion of the arguments. □

The result stated above establishes the necessary and sufficient condition for matrices of $\mathbb{C}^{k\times 2}_n$,$k\geq 4$ to be $\mathcal{E}^c_{eh}$-equivalent. The property expressed by eqn(6.28) will be referred to in short as <u>collinearity</u> of the r-prime Plücker vector sets $\mathcal{P}^1_r,\mathcal{P}^2_r$ of T_1,T_2 respectively. Some interesting remarks and a corollary are stated next.

<u>Remark (6.8)</u>: The collinearity property of $\mathcal{P}^1_r,\mathcal{P}^2_r$ implies collinearity of $\mathcal{P}^1_j,\mathcal{P}^2_j$ for any other $j\in\langle k\rangle$. Thus for testing $\mathcal{E}^c_{eh}$-equivalence of T_1,T_2, it is enough to check the collinearity of $\mathcal{P}^1_1,\mathcal{P}^2_1$.

<u>Corollary (6.2)</u>: The set of Plücker vectors $\mathcal{P}_r$ which correspond to any fixed index $r\in\langle k\rangle$, forms a complete and independent set of invariants for the following cases:

(i) Matrices $T\in\mathbb{C}^{k\times 2}_n$,$k\geq 4$, under $\mathcal{E}^c_{eh}$-equivalence.

(ii) Matrices $T\in\mathbb{R}^{k\times 2}_n$,$k\geq 4$, under $\mathcal{E}^r_{eh}$-equivalence. □

Note that the independence of the set $\mathcal{P}_r$ follows from the proof of Theorem (6.3); that is, the set $\mathcal{P}_r$, or equivalently the set of conditions (6.32c) is the minimal set of conditions needed for (6.32b) to be true for all $(1,i_1,i_2)\in Q^{1]}_{3,k}$. The set of parameters ξ_i and the parameter δ are defined by the following result.

<u>Corollary (6.3)</u>: Let $T_1,T_2\in\mathbb{C}^{k\times 2}_n$ $(\mathbb{R}^{k\times 2}_n)$,$k\geq 4$, and let the corresponding 1-prime Plücker vectors $\mathcal{P}^1_1,\mathcal{P}^2_1$ be collinear. The scaling parameters ξ_i and the determinant δ of the transformation Q for which $T_1\mathcal{E}^c_{eh}T_2$ $(T_1\mathcal{E}^r_{eh}T_2)$ are

defined by:

(i) $\delta\xi_1^2$ is given as the common ratio

$$\frac{a^2_{1,i_1}}{a^1_{1,i_1}}\cdot\frac{a^2_{1,i_2}}{a^1_{1,i_2}}\cdot\frac{a^1_{i_1,i_2}}{a^2_{i_1,i_2}} = \delta\xi_1^2 = c \text{ for } \forall\ (1,i_1,i_2)\in Q^{1]}_{3,k} \tag{6.33a}$$

(ii) Using the value of $\delta\xi_1$ defined by (6.33a), then

$$\xi_i = a^2_{1,i}/a^1_{1,i}\,\delta\xi_1 \text{ for } \forall\ i=2,3,\ldots k \tag{6.33b}$$

□

With the choice of $\delta,\xi_i,i\in\underset{\sim}{k}$ as above, the matrices T_2 and $\text{diag}\{\xi_i\}T_1$ have the same column space (since their Grassmann vectors differ by a scalar) and thus Q may be defined as the coordinate transformation ($Q\in\mathbb{C}^{2\times 2}$, $|Q|=\delta\in\mathbb{R}-\{0\}$) relating the two bases of the same space. Finally, we note the following:

Remark (6.9): The set of all Plücker vectors $\mathcal{P}$ which correspond to all indices $i\in\underset{\sim}{k}$, $\mathcal{P}\triangleq\mathcal{P}_1\cup\mathcal{P}_2\cup\ldots\cup\mathcal{P}_k$, is a complete, but dependent set of invariants of $T\in\mathbb{C}_n^{k\times 2}$ ($T\in\mathbb{R}_n^{k\times 2}$), $k\geq 4$ under $\mathcal{E}^c_{eh}$ ($\mathcal{E}^r_{eh}$) equivalence. Any subset $\mathcal{P}_j$ of $\mathcal{P}$, which corresponds to only one index $r\in\langle k\rangle$, provides a complete and independent set of invariants.

6.5.3 A complete set of invariants for matrices of $\mathbb{C}_n^{k\times 2}$ under $\mathcal{E}^r_{eh}$-equivalence

The task we have originally set was the investigation of the conditions under which two matrices $T_1,T_2\in\mathbb{C}_n^{k\times 2}$ may be $\mathcal{E}^r_{eh}$ equivalent. For $T_1,T_2\in\mathbb{R}_n^{k\times 2}$, Proposition (6.11) and Remark (6.7) show that if $k\leq 3$, then we always have that $T_1\mathcal{E}^r_{eh}T_2$; if $k\geq 4$, the Corollary (6.2) shows that the set of r-prime Plücker vectors $\mathcal{P}_r$ is a complete and independent set of invariants for $\mathcal{E}^r_{eh}$-equivalence of matrices of $\mathbb{R}_n^{k\times 2}$. The general case of $\mathcal{E}^r_{eh}$-equivalence of matrices of $\mathbb{C}_n^{k\times 2}$ is considered here. We first note the following.

Corollary (6.4): The set of r-prime Plücker vectors P_r, $r\in\langle k\rangle$ fixed, defines an independent, but not complete set of invariants for matrices of $\mathbb{C}_n^{k\times 2}$, $k\geq 4$, under E^r_{eh}-equivalence. □

The above result is a mere consequence of the fact that a complete and independent set of invariants for matrices of $\mathbb{C}_n^{k\times 2}$, $k\geq 4$, under E^c_{eh}-equivalence, defines necessary conditions for E^r_{eh}-equivalence, but not sufficient. Our task, therefore, is to define here the extra conditions needed for the transformation Q with a real determinant to be also real ($Q\in\mathbb{R}^{2\times 2}$). Such a study yields extra conditions and thus an additional set of invariants for E^r_{eh}-equivalence. Our study is greatly simplified by adopting the following notation.

Definition (6.4):

(i) Let $\xi=v+jw\in\mathbb{C}$, $v,w\in\mathbb{R}$. Then, a real matrix representation of ξ, $[\xi]_{\mathbb{R}}$, is defined by

$$[\xi]_{\mathbb{R}} \triangleq \begin{bmatrix} v & -w \\ w & v \end{bmatrix} \in \mathbb{R}^{2\times 2} \tag{6.34}$$

(ii) Let $\mathbb{C}\times\mathbb{C}=\{e:e=(\sigma+j\omega,-\sigma'-j\omega'),\sigma,\sigma',\omega,\omega'\in\mathbb{R}\}$. The operation $[\ ]_{\mathbb{R}}:\mathbb{C}\times\mathbb{C}\rightarrow\mathbb{R}^{2\times 2}$ is defined for $\forall\ e=(\sigma+j\omega,-\sigma'-j\omega')\in\mathbb{C}\times\mathbb{C}$ by

$$[e]_{\mathbb{R}} \triangleq \begin{bmatrix} \sigma & -\sigma' \\ \omega & -\omega' \end{bmatrix} \in \mathbb{R}^{2\times 2} \tag{6.35}$$

$[e]_{\mathbb{R}}$ will be referred to as the real matrix representation of e. Every $e=(\sigma+j\omega,-\sigma'-j\omega')\in\mathbb{C}\times\mathbb{C}$ defines a row vector $\underline{e}^t=[\sigma+j\omega,-\sigma'-j\omega']\in\mathbb{C}^{1\times 2}$; the operation $[\cdot]_{\mathbb{R}}$ on $\underline{e}^t$ is defined in the same way as for the ordered pair (eqn(6.35)).

(iii) Let $T=[\dots,\underline{e}_i,\dots]^t\in\mathbb{C}^{k\times 2}$, where $\underline{e}_i^t=[\sigma_i+j\omega_i,-\sigma_i'-j\omega_i']$. The real matrix representation of T is defined by

$$[\,T\,]_{\mathbb{R}} \triangleq \begin{bmatrix} \vdots \\ [\underline{e}_i^t]_{\mathbb{R}} \\ \vdots \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \hdashline \sigma_i & & -\sigma_i' \\ \omega_i & & -\omega_i' \\ \hdashline & \vdots & \end{bmatrix} \in \mathbb{R}^{2k\times 2},\ i\in\underset{\sim}{k} \tag{6.36}$$

□

Within this notation, it may be readily verified that for $\forall\ \xi=v+jw\in\mathbb{C}$ and $e=(\sigma+j\omega,-\sigma'-j\omega')\in\mathbb{C}\times\mathbb{C}$, then

$$[\xi.e]_{\mathbb{R}} = [\,\xi\,]_{\mathbb{R}}[e\,]_{\mathbb{R}} = \begin{bmatrix} v & -w \\ w & v \end{bmatrix}\begin{bmatrix} \sigma & -\sigma' \\ \omega & -\omega' \end{bmatrix} \tag{6.37}$$

Using the above property we may express E^r_{eh}-equivalence as follows: $T_1 E^r_{eh} T_2$, iff there exist $\xi_i\in\mathbb{C}-\{0\}$, $i\in\underset{\sim}{k}$, and a $Q\in\mathbb{R}^{2\times 2}$, $|Q|\neq 0$, such that

$$[T_2]_{\mathbb{R}} = \mathrm{diag}\{\ldots,[\xi_i]_{\mathbb{R}},\ldots\}\,[T_1]_{\mathbb{R}}\,Q \tag{6.38}$$

Note that (6.38) is a <u>real representation</u> of E^r_{eh}-equivalence and thus provides an equivalent definition for E^r_{eh}-equivalence of matrices of $\mathbb{C}_n^{k\times 2}$. This equivalent definition will be used in the following, for defining the extra conditions that guarantee a real Q.

<u>Proposition (6.12)</u>: Let $T_1,T_2\in\mathbb{C}_n^{k\times 2}$ and let $\xi_i=x_i+jy_i\in\mathbb{C}-\{0\}$ with $|\delta|=1$ be a solution of (6.26a). Necessary and sufficient condition for $T_1 E^r_{eh} T_2$ is that either of the following conditions hold true:

(i) $$\underline{g}([T_2]_{\mathbb{R}}) = \varepsilon\,\underline{g}(\mathrm{diag}([\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}),\quad \varepsilon=\pm 1 \tag{6.39a}$$

or

(ii) $$\underline{g}([T_2]_{\mathbb{R}}) = \varepsilon\,\underline{g}(\bar{E}\ \mathrm{diag}([\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}),\quad \varepsilon=\pm 1 \tag{6.39b}$$

where $\bar{E}=\mathrm{diag}\{E,\ldots,E\}$ and $E=[j]_{\mathbb{R}}$.

<u>Proof</u>

By Proposition (6.10), if $\xi_i\in\mathbb{C}-\{0\}$, $i\in\underset{\sim}{k}$, is a solution of (6.26a) for $|\delta|=1$, then any other solution with $|\delta|=1$ is defined by $\xi_i'=j\xi_i$, $i\in\underset{\sim}{k}$. For a real Q to exist either of the following conditions must hold true:

$$[T_2]_{\mathbb{R}} = \mathrm{diag}([\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}\,Q \tag{6.40a}$$

or

$$[T_2]_{\mathbb{R}} = \mathrm{diag}([j\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}\,Q \tag{6.40b}$$

Note that $[j\xi_i]_{\mathbb{R}}=[j]_{\mathbb{R}}[\xi_i]_{\mathbb{R}}=E[\xi_i]_{\mathbb{R}}$ and thus (6.40b) is equivalent to

$$[T_2]_{\mathbb{R}} = \mathrm{diag}(E,\dots,E)\mathrm{diag}([\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}\,Q \tag{6.40c}$$

Necessary and sufficient condition for (6.40a), or (6.40c) to have a solution for a matrix Q is that

$$\text{col-span}_{\mathbb{R}}\{[T_2]_{\mathbb{R}}\} = \text{col-span}_{\mathbb{R}}\{\mathrm{diag}([\xi_i]_{\mathbb{R}})[T_1]_{\mathbb{R}}\} \tag{6.40d}$$

or

$$\text{col-span}_{\mathbb{R}}\{[T_2]_{\mathbb{R}}\} = \text{col-span}_{\mathbb{R}}\{\mathrm{diag}(E[\xi_i]_{\mathbb{R}})[T_1]_{\mathbb{R}}\} \tag{6.40e}$$

If either (6.40d), or (6.40e), holds true, then either $T_2=\mathrm{diag}\{\xi_i\}T_1Q$, or $T_2=\mathrm{diag}\{j\xi_i\}T_1Q$. However, because the ξ_i's have been chosen with $|\delta|=1$, then $|Q|=\pm1=\varepsilon$. Conditions (6.40d), or (6.40e), may then be translated in terms of the Grassmann vectors, by conditions (6.39a), or (6.39b), respectively. □

By combining now Propositions (6.12), (6.9) and (6.10) we obtain the necessary and sufficient conditions for matrices of $\mathbb{C}_n^{k\times2}$ to be E_{eh}^r-equivalent.

Proposition (6.13): Let $T_1,T_2\epsilon\mathbb{C}_n^{k\times2}$, $k\geq4$, and let P_r^1,P_r^2 be the r-prime Plücker vectors of T_1,T_2 respectively defined for a fixed $r\epsilon<k>$. Necessary and sufficient conditions for $T_1E_{eh}^rT_2$ are:

(i) The sets P_r^1 and P_r^2 are collinear.

(ii) If ξ_i is a solution of (6.26a), with $|\delta|=1$, as defined by Corollary (6.3), then at least one of the following two conditions hold true

$$\underline{g}([T_2]_{\mathbb{R}}) = \varepsilon\, \underline{g}(\mathrm{diag}([\xi_i]_{\mathbb{R}})\,[T_1]_{\mathbb{R}}),\quad \varepsilon=\pm 1 \tag{6.41a}$$

or

$$\underline{g}([T_2]_{\mathbb{R}}) = \varepsilon\, \underline{g}(\mathrm{diag}(E[\xi_i]_{\mathbb{R}})[T_1]_{\mathbb{R}}),\quad \varepsilon=\pm 1 \tag{6.41b}$$

where $E=[j]_{\mathbb{R}}$ and $\varepsilon=\pm 1$. □

Note that condition (i) guarantees the existence of the ξ_i for E^c_{eh}-equivalence, whereas conditions (ii) guarantee that the transformation Q is real. For the case of matrices of $\mathbb{C}^{k\times 2}_n$ with k=3 we have the following result.

Corollary (6.5): Let $T_1,T_2\in\mathbb{C}^{3\times 2}_n$, and let $\underline{g}(T_1)=[a^1_{1,2},a^1_{1,3},a^1_{2,3}]^t$, $\underline{g}(T_2)=[a^2_{1,2},a^2_{1,3},a^2_{2,3}]^t$ and let $D=a^2_{1,2}a^2_{1,3}a^2_{2,3}/a^1_{1,2}a^1_{1,3}a^1_{2,3}$, $\hat{q}_{1,2}=a^1_{1,2}/a^2_{1,2}$, $\hat{q}_{1,3}=a^1_{1,3}/a^2_{1,3}$, $\hat{q}_{2,3}=a^1_{2,3}/a^2_{2,3}$. Necessary and sufficient condition for $T_1\,{}^{r}_{eh}\,T_2$ is that either of the following conditions hold true:

(i) $\underline{g}([T_2]_{\mathbb{R}}) = \lambda \underline{g}([\sqrt{D}\,\mathrm{diag}\{\hat{q}_{2,3},\hat{q}_{1,3},\hat{q}_{1,2}\}]_{\mathbb{R}}\,[T_1]_{\mathbb{R}}),\quad \lambda\in\mathbb{R}-\{0\}$ (6.42a)

or

(ii) $\underline{g}([T_2]_{\mathbb{R}})=\lambda'\underline{g}(\bar{E}[\sqrt{D}\,\mathrm{diag}\{\hat{q}_{2,3},\hat{q}_{1,3},\hat{q}_{1,2}\}]_{\mathbb{R}}\,[T_1]_{\mathbb{R}}),\quad \lambda'\in\mathbb{R}-\{0\}$ (6.42b)

where $\bar{E}=\mathrm{diag}(E,E,E)$ and $E=[j]_{\mathbb{R}}$.

The proof of the above result follows by a mere substitution of the ξ_i's, as they are given by (6.27a) and (6.27b), and by Proposition (6.12) (note now that $|\delta|$ is not assumed to be 1). Before we examine the cases k=1,2 we give an example to illustrate the case of k≥4.

Example (6.2): Let $T_1,T_2\in\mathbb{C}^{4\times 2}_n$, where

$$T_1 = \begin{bmatrix} 1+j & 1-j \\ 1 & 2+j \\ 2-j & 3 \\ 1 & 1+2j \end{bmatrix},\quad T_2 = \begin{bmatrix} 3+j & 2 \\ 9-2j & 7-j \\ 14-4j & 10-2j \\ 1+5j & 4j \end{bmatrix}$$

In order to check E^r_{eh}-equivalence, we compute first the Grassmann vectors, i.e.

$$\underline{g}(T_1)=[a^1_{1,2},a^1_{1,3},a^1_{1,4},a^1_{2,3},a^1_{2,4},a^1_{3,4}]^t=[4j,2+6j,-2+4j,-2,-1+j,1+3j]^t$$

$$\underline{g}(T_2)=[a^2_{1,2},a^2_{1,3},a^2_{1,4},a^2_{2,3},a^2_{2,4},a^2_{3,4}]^t=[4+8j,4+12j,-6+2j,-8+4j,-4+2j,-4+8j]^t$$

There is only one prime quadruple , $\phi=(1,2,3,4)$, one prime triangle $\delta\phi$, $\delta\phi=\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$ and thus each of T_1,T_2 has one Plücker vector. These are

$$\underline{p}^1(\delta\phi) = \begin{bmatrix} a^1_{1,2}a^1_{3,4} \\ -a^1_{1,3}a^1_{2,4} \\ a^1_{1,4}a^1_{2,3} \end{bmatrix} = \begin{bmatrix} -12+4j \\ 8+4j \\ 4-8j \end{bmatrix} , \ \underline{p}^2(\delta\phi) = \begin{bmatrix} a^2_{1,2}a^2_{3,4} \\ -a^2_{1,3}a^2_{2,4} \\ a^2_{1,4}a^2_{2,3} \end{bmatrix} = \begin{bmatrix} -80 \\ 40+40j \\ 40-40j \end{bmatrix}$$

We can readily see that $\underline{p}^2(\delta\phi)=(6+2j)\underline{p}^1(\delta\phi)$; thus $T_1 E^c_{eh} T_2$ and the set of ξ_i, $i\in\underset{\sim}{4}$ and that a $\delta\in\mathbb{R}-\{0\}$ exist. If we take $\delta=1$, then by Corollary (6.3) we have that $\xi_1=1$, $\xi_2=2-j$, $\xi_3=2$, $\xi_4=1+j$. To check the existence of a real Q we compute $[T_2]_{\mathbb{R}}$, $[\mathrm{diag}\{\xi_i\}]_{\mathbb{R}}[T_1]_{\mathbb{R}}$ and then the corresponding Grassmann vectors. Thus,

$$[T_2]_{\mathbb{R}} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \\ 9 & 7 \\ -2 & -1 \\ 14 & 10 \\ -4 & -2 \\ 1 & 0 \\ 5 & 4 \end{bmatrix} , \ [\mathrm{diag}\{\xi_i\}T_1]_{\mathbb{R}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 5 \\ -1 & 0 \\ 4 & 6 \\ -2 & 0 \\ 1 & -1 \\ 1 & 3 \end{bmatrix}$$

and it may be readily verified that $\underline{g}([T_2]_{\mathbb{R}})=(1)\underline{g}([\mathrm{diag}\{\xi_i\}T_1]_{\mathbb{R}})$; the latter implies the existence of a $Q\in\mathbb{R}^{2\times 2}$ with $\delta=|Q|=1$ for which $T_1 E^r_{eh} T_2$. The corresponding Q is

$$Q = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } |Q| = \delta = 1$$

□

Proposition (6.13) gives the necessary and sufficient conditions for $T_1,T_2\in\mathbb{C}_n^{k\times 2}$ to be $\mathcal{E}^r_{eh}$-equivalent. Note, however, that the set of conditions (ii) of the result depend upon the parameters ξ_i, the existence of which is guaranteed by part (i) of the theorem. Working out necessary and sufficient conditions which are independent from the solution parameters ξ_i, is equivalent to the problem of finding extra invariants for $\mathcal{E}^r_{eh}$-equivalence, which together with the Plücker vectors form an independent and complete set. Before we proceed to the definition of the extra invariants we note:

<u>Remark (6.10)</u>: If the pair $p=(\sigma+j\omega,-\sigma'-j\omega')$ represents a point $P\in\mathbb{C}-\mathbb{R}$ then

$$\det[P]_{\mathbb{R}} = \begin{bmatrix} \sigma & -\sigma' \\ \omega & -\omega' \end{bmatrix} \neq 0$$

A matrix $T\in\mathbb{C}_n^{k\times 2}$ for which the ordered pair of the elements of every row represents a point in $\mathbb{C}-\mathbb{R}$, will be referred to as <u>purely complex and entirely nonsingular matrix</u> and the set of such matrices will be denoted by $\bar{\mathbb{C}}_n^{k\times 2}$. Note that matrices in $\mathbb{C}_n^{k\times 2}$ may have a row representing a point in $\mathbb{R}$.

The search for some extra invariants of matrices of $T\in\mathbb{C}_n^{k\times 2}$ under $\mathcal{E}^r_{eh}$-equivalence is initiated by considering the simple cases k=1,2.

<u>Proposition (6.14)</u>: Let $T_1,T_2\in\mathbb{C}_n^{k\times 2}$. Then $T_1\mathcal{E}^r_{eh}T_2$ is always true in the following cases:

(i) $T_1,T_2\in\mathbb{C}^{1\times 2}$

(ii) $T_1,T_2\in\mathbb{R}_n^{k\times 2}$, where k=1,2,3.

(iii) $T_1,T_2\in\mathbb{C}_n^{2\times 2}$, where

$$T_1 = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix},\ T_2 = \begin{bmatrix} \tilde{\alpha}_1 & \tilde{\beta}_1 \\ \tilde{\alpha}_2 & \tilde{\beta}_2 \end{bmatrix} \text{ and } \alpha_1,\beta_1,\tilde{\alpha}_1,\tilde{\beta}_1\in\mathbb{R},\ \alpha_2,\beta_2,\tilde{\alpha}_2,\tilde{\beta}_2\in\mathbb{C}$$

Proof

The cases (i),(ii) have been already proved before.

(iii) $T_1 \mathcal{E}^r_{eh} T_2$ implies that $\xi_1 \in \mathbb{R}, \xi_2 \in \mathbb{C}$ and $Q \in \mathbb{R}^{2\times 2}$ such that

$$\xi_1[\alpha_1,\beta_1]Q = [\tilde{\alpha}_1,\tilde{\beta}_1],\ \xi_2[\alpha_2,\beta_2]Q = [\tilde{\alpha}_2,\tilde{\beta}_2] \tag{6.43a}$$

and by taking real representations we have

$$\xi_1[\alpha_1,\beta_1]Q = [\tilde{\alpha}_1,\tilde{\beta}_1],\ \begin{bmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{bmatrix}\begin{bmatrix} \sigma_1 & -\sigma_1' \\ \omega_1 & -\omega_1' \end{bmatrix} Q = \begin{bmatrix} \sigma_2 & -\sigma_2' \\ \omega_2 & -\omega_2' \end{bmatrix} \tag{6.43b}$$

By solving the second of (6.43b) and substituting into the first and then by rearranging the terms we have

$$\underbrace{\xi_1[\alpha_1,\beta_1]\begin{bmatrix} \sigma_1 & -\sigma_1' \\ \omega_1 & -\omega_1' \end{bmatrix}^{-1}}_{\triangleq\ [k_1,m_1]} = \underbrace{[\tilde{\alpha}_1,\tilde{\beta}_1]\begin{bmatrix} \sigma_2 & -\sigma_2' \\ \omega_2 & -\omega_2' \end{bmatrix}^{-1}}_{\triangleq\ [k_2,m_2]}\begin{bmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{bmatrix} \tag{6.43c}$$

or equivalently

$$[k_1,m_1] = [k_2,m_2]\begin{bmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{bmatrix} \Longleftrightarrow \begin{cases} k_2x_2+m_2y_2 = k_1 \\ m_2x_2-k_2y_2 = m_1 \end{cases} \tag{6.43d}$$

Note that $(k_1,m_1),(k_2,m_2)\neq(0,0)$ and that (6.43d) has a solution if and only if $k_2^2+m_2^2\neq 0 \iff (k_2,m_2)\neq(0,0)$; thus, by assuming $\xi_1 \in \mathbb{R}-\{0\}$, arbitrary, we can find $\xi_2=x_2+jy_2$ by solving (6.43d). The matrix Q is then given by

$$Q = \begin{bmatrix} \sigma_1 & -\sigma_1' \\ \omega_1 & -\omega_1' \end{bmatrix}^{-1}\begin{bmatrix} x_2 & -y_2 \\ y_2 & x_2 \end{bmatrix}^{-1}\begin{bmatrix} \sigma_2 & -\sigma_2' \\ \omega_2 & -\omega_2' \end{bmatrix} \tag{6.43e}$$

□

Proposition (6.15): Let $T_1,T_2 \in \bar{\mathbb{C}}_n^{2\times 2}$, $[T_i]_{\mathbb{R}} = [\Sigma_1^{it},\Sigma_2^{it}]^t$, i=1,2, and let

$$M = \left[\begin{array}{cc:cc} a_2 & b_2 & a_1 & c_1 \\ b_2 & -a_2 & b_1 & d_1 \\ c_2 & d_2 & c_1 & -a_1 \\ d_2 & -c_2 & d_1 & -b_1 \end{array}\right] = \begin{bmatrix} P_2^{\alpha} & P_1^{\beta} \end{bmatrix} \tag{6.44a}$$

where the parameters in M are defined by

$$A_i = \Sigma_2^i(\Sigma_1^i)^{-1} = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}, \ i=1,2 \tag{6.44b}$$

$T_1 E^r_{eh} T_2$, if and only if $|M|=0$.

<u>Proof</u>

$T_1 E^r_{eh} T_2$ implies that $\xi_i=x_i+jy_i \in \mathbb{C}-\{0\}$, $i=1,2$ and a $Q \in \mathbb{R}^{2\times 2}$, $|Q| \neq 0$ such that if $Z_i=[\xi_i]_{\mathbb{R}}$, $i=1,2$, then

$$\Sigma_i^2 = Z_i \Sigma_i^1 Q, \ i=1,2 \tag{6.45a}$$

By solving the first of (6.45a) and substituting into the second we have

$$\Sigma_2^2(\Sigma_1^2)^{-1}Z_1 = Z_2\Sigma_2^1(\Sigma_1^1)^{-1} \Longleftrightarrow A_2Z_1 = Z_2A_1 \tag{6.45b}$$

If we define by $\underline{q}=[x_1,y_1,-x_2,y_2]^t$, then (6.45b) may be equivalently expressed as

$$M\underline{q} = 0 \tag{6.45c}$$

Clearly, a nonzero vector $\underline{q}$ exists if and only if $|M|=0$. Note that any vector $\underline{q} \in N_r(M)$ has the property that $(x_i,y_i) \neq (0,0)$, $i=1,2$. Indeed, assume that $x_1=y_1=0$, then (6.45c) yields

$$\begin{bmatrix} a_1 & c_1 \\ b_1 & d_1 \end{bmatrix} \begin{bmatrix} -x_2 \\ y_2 \end{bmatrix} = 0 \tag{6.45d}$$

which however contradicts the full rank assumption of A_1 $(T_1, T_2 \in \bar{\mathbb{C}}_n^{2\times 2})$. For similar reasons, vectors in $N_r(M)$ with $x_2=y_2=0$ do not exist. The sufficiency is eastablished by a mere reversion of the steps. □

From (6.44a) and (6.44b) it is clear that for a matrix $T_i \in \bar{\mathbb{C}}_n^{2\times 2}$ we may define two matrices P_i^α, $P_i^\beta \in \mathbb{R}^{4\times 2}$; P_i^α, P_i^β will be called respectively the <u>α-, β-characteristic matrices</u> of T_i. If we denote the property that

$|[P_2^{\alpha}, P_1^{\beta}]|=0$ by $P_2^{\alpha} \Box P_1^{\beta}=0$ then the conditions of Proposition (6.15) may be expressed by

$$P_2^{\alpha} \Box P_1^{\beta} = P_1^{\beta} \Box P_2^{\alpha} = 0 \iff P_1^{\alpha} \Box P_2^{\beta} = P_2^{\beta} \Box P_1^{\alpha} = 0 \tag{6.46}$$

From the above results we can state the following remark.

Remark (6.11): There always exists a $\beta \in PGL(1,\mathbb{C}/\mathbb{R})$ that maps:

(i) A set of one, two, or three real distinct points to any other set of one, two, or three real distinct points of $\mathbb{C} \cup \{\infty\}$ respectively.

(ii) A set made up from a real point and a pair of complex conjugate points to any other set made up from a real point and a pair of complex conjugate points of $\mathbb{C} \cup \{\infty\}$.

Note that two distinct complex points of the $\mathbb{C}$-plane may not always be mapped by a $\beta \in PGL(1,\mathbb{C}/\mathbb{R})$ to any two distinct points of the $\mathbb{C}$-plane. The condition under which such a mapping exists is defined by the rank deficiency of the matrix M.

Our attention is focussed next on the cases where $k \geq 3$. The main idea is to substitute the expressions of the ξ_i's, as defined by Corollary (6.3), into the conditions (ii) of Theorem (6.4) and thus derive conditions for the existence of a real Q, which are independent of the ξ_i's. For the case of k=3, Corollary (6.5) provides such conditions; however, some further rearrangement is needed for the definition of new invariants.

The expression of the ξ_i's is given by Corollary (6.3). Note that the solution parameters in (6.33a) and (6.33b) have been expressed in terms of ξ_1, δ and some indices $(i_1,i_2) \in Q_{2,k-1}^{1)}$. Such a parametrization of the solution will be referred to as <u>$(1,i_1,i_2)$-parametrization</u> of the E_{eh}^{c}-equivalence parameters. Clearly, the solution parameters may be parametrized in terms of any other ξ_r, $r \in \langle k \rangle$ and any $(j_1,j_2) \in Q_{2,k-1}^{r)}$, thus leading to an equivalent (r,j_1,j_2)-parametrization. For the sake of simplicity in the following we shall use the $(1,i_1,i_2)$-parametrization.

<u>Definition (6.5)</u>: Let $T\in\mathbb{C}_n^{k\times 2}$ or $T\in\mathbb{R}_n^{k\times 2}$, $k\geqslant 3$ and let $a_\omega, \omega=(j_1,j_2)\in Q_{2,k}$ be the coordinates of $\underline{g}(T)$.

(i) We define as the <u>$(1,i_1,i_2)$-standardising matrix of T</u>, the matrix

$$S_{1,i_1,i_2} = c_{1,i_1,i_2}\ \mathrm{diag}\{a_{i_1,i_2}/a_{1,i_1}a_{1,i_2}, 1/a_{1,2}, \dots, 1/a_{1,k}\} \tag{6.47a}$$

where $(i_1,i_2)\in Q_{2,k}^{1)}$ and

$$c_{1,i_1,i_2} = (a_{1,i_1}a_{1,i_2}/a_{i_1,i_2})^{1/2},\ \text{if } T\in\mathbb{C}_n^{k\times 2} \tag{6.47b}$$

$$c_{1,i_1,i_2} = (|a_{1,i_1}a_{1,i_2}/a_{i_1,i_2}|)^{1/2},\ \text{if } T\in\mathbb{R}_n^{k\times 2} \tag{6.47c}$$

(ii) If $T\in\mathbb{C}_n^{k\times 2}$, then we define as the <u>$(1,i_1,i_2)$-$\mathbb{C}$-standard form</u> T_{1,i_1,i_2}, and as the <u>$(1,i_1,i_2)$-$\mathbb{R}$-standard form</u>, F_{1,i_1,i_2}, of T the matrices

$$T_{1,i_1,i_2} \triangleq S_{1,i_1,i_2}T\in\mathbb{C}_n^{k\times 2},\ F_{1,i_1,i_2} \triangleq [T_{1,i_1,i_2}]_{\mathbb{R}} \in\mathbb{R}^{2k\times 2} \tag{6.48}$$

If $T\in\mathbb{R}_n^{k\times 2}$, then the <u>$(1,i_1,i_2)$-$\mathbb{R}$-standard form</u> of T is defined by the matrix $T_{1,i_1,i_2}\in\mathbb{R}_n^{k\times 2}$ given by (6.48).

(iii) For a matrix $T\in\mathbb{C}_n^{k\times 2}$, the Grassmann vectors $\underline{g}(T_{1,i_1,i_2})\triangleq \underline{g}_{1,i_1,i_2}$ and $\underline{g}(F_{1,i_1,i_2})\triangleq\tilde{\underline{g}}_{1,i_1,i_2}$ are defined as the <u>$(1,i_1,i_2)$-$\mathbb{C}$-canonical Grassmann vector</u> ($(1,i_1,i_2)$-$\mathbb{C}$-CGV) and <u>$(1,i_1,i_2)$-$\mathbb{R}$-canonical Grassmann vector</u> ($(1,i_1,i_2)$-$\mathbb{R}$-CGV) of T respectively. Similarly, if $T\in\mathbb{R}_n^{k\times 2}$, then $\underline{g}(T_{1,i_1,i_2})\triangleq\tilde{\underline{g}}_{1,i_1,i_2}$ is defined as the <u>$(1,i_1,i_2)$-$\mathbb{R}$-canonical Grassmann vector</u> ($(1,i_1,i_2)$-$\mathbb{R}$-CGV) of T. □

Clearly, similar definitions may be given with respect to any triple $(r,i_1,i_2), r\in\langle k\rangle, (i_1,i_2)\in Q_{2,k-1}^{r)}$; the corresponding $\underline{g}_{r,i_1,i_2}, \tilde{\underline{g}}_{r,i_1,i_2}$ will be referred to in short as (r,i_1,i_2)-$\mathbb{C}$-CGV, (r,i_1,i_2)-$\mathbb{R}$-CGV respectively. Note that $T_{r,i_1,i_2}\mathcal{E}_{eh}^{c}T$ and $T_{r,i_1,i_2}\mathcal{E}_{eh}^{r}T$. The importance of the T_{1,i_1,i_2} F_{1,i_1,i_2} matrices and of the corresponding $\underline{g}_{1,i_1,i_2}, \tilde{\underline{g}}_{1,i_1,i_2}$ for $\mathcal{E}_{eh}^{c}, \mathcal{E}_{eh}^{r}$ equivalence respectively is discussed next.

Theorem (6.4): Let $T_1,T_2\in\mathbb{C}_n^{k\times 2}$, k≥3 and let $g^1_{1,i_1,i_2}, g^2_{1,i_1,i_2}$ be the corresponding $(1,i_1,i_2)$-ℂ-CGV of T_1,T_2 respectively. Necessary and sufficient condition for $T_1 E^c_{eh} T_2$ is that

$$g^1_{1,i_1,i_2} = g^2_{1,i_1,i_2} \tag{6.49}$$

Proof

Let $T_1 E^c_{eh} T_2$ and let ξ_i be the set of parameters expressed by Corollary (6.3) in the $(1,i_1,i_2)$ parametric form. By substituting the ξ_i's in $T_2=\mathrm{diag}\{\xi_i\}T_1Q$ we have that

$$T_2 = \frac{1}{\delta\xi_1}\,\mathrm{diag}\{\delta\xi_1^2, \frac{a^2_{1,2}}{a^1_{1,2}},\dots,\frac{a^2_{1,k}}{a^1_{1,k}}\}T_1Q,\quad Q\in\mathbb{C}_n^{2\times 2},\quad |Q|=\delta\in\mathbb{R} \tag{6.50a}$$

where $\delta\xi_1^2$ is given by (6.33a) and thus

$$1/\delta\xi_1 = \varepsilon c^1_{1,i_1,i_2}/c^2_{1,i_1,i_2}\sqrt{\delta},\quad \varepsilon=\pm1 \tag{6.50b}$$

where c^j_{1,i_1,i_2}, j=1,2 is the standardising parameter of the T_j matrix. By (6.50a) and (6.50b), it follows that

$$c^2_{1,i_1,i_2}T_2 = \frac{\varepsilon}{\sqrt{\delta}}\,c^1_{1,i_1,i_2}\quad \mathrm{diag}\{\frac{a^1_{i_1,i_2}}{a^1_{1,i_1}a^1_{1,i_2}}\cdot\frac{a^2_{1,i_1}a^2_{1,i_2}}{a^2_{i_1,i_2}}, \frac{a^2_{1,2}}{a^1_{1,2}},\dots,\frac{a^2_{1,k}}{a^1_{1,k}}\}T_1Q$$

and thus

$$T^2_{1,i_1,i_2} = \frac{\varepsilon}{\sqrt{\delta}}\,T^1_{1,i_1,i_2}Q \tag{6.50c}$$

By applying the Binet-Cauchy Theorem on (6.50c) it follows that

$$g^2_{1,i_1,i_2} = C_2\{\frac{\varepsilon}{\sqrt{\delta}}T^1_{1,i_1,i_2}Q\} = \frac{1}{\delta}\,g^1_{1,i_1,i_2}|Q| = g^1_{1,i_1,i_2} \tag{6.50d}$$

and this proves the necessity. To prove the sufficiency we start from $g^2_{1,i_1,i_2}=g^1_{1,i_1,i_2}$. Then, given that by construction $g^2_{1,i_1,i_2}, g^1_{1,i_1,i_2}$ are decomposable, (6.50d) implies that there exists a $Q'\in\mathbb{C}_n^{2\times 2}$ with $|Q'|=1$ such that $T^2_{1,i_1,i_2} = T^1_{1,i_1,i_2} Q'$, or that $T^2_{1,i_1,i_2} \mathcal{E}^c_{eh} T^1_{1,i_1,i_2}$. Given that $T^2 \mathcal{E}^c_{eh} T^2_{1,i_1,i_2}$ and $T^1 \mathcal{E}^c_{eh} T^1_{1,i_1,i_2}$ it follows that $T_1 \mathcal{E}^c_{eh} T_2$. □

Corollary (6.6): The $(1,i_1,i_2)$-$\mathbb{C}$-CGV $\underline{g}_{1,i_1,i_2}$ is a complete invariant for matrices of $\mathbb{C}^{k\times 2}_n$, $k\geq 3$, under $\mathcal{E}^c_{eh}$-equivalence. □

Remark (6.12): For all matrices $T\in\mathbb{C}^{3\times 2}_n$ the (1,2,3)-CGV $\underline{g}_{1,2,3}$ is given by $\underline{g}_{1,2,3}=[1,1,1]^t$ and thus all matrices of $\mathbb{C}^{3\times 2}_n$ are $\mathcal{E}^c_{eh}$-equivalent. This provides an alternative proof for Proposition (6.11). For such class of matrices $\underline{g}_{1,2,3}$ is not an essential invariant, since all elements of the set possess this property.

Remark (6.13): The $(1,i_1,i_2)$-$\mathbb{C}$-CGV $\underline{g}_{1,i_1,i_2}$, or any $\underline{g}_{r,i_1,i_2}$ vector, is an equivalent complete invariant for $\mathcal{E}^c_{eh}$-equivalence to that defined by the set of r-prime Plücker vectors $\mathcal{P}_r$.

Theorem (6.5) applies also to the case of $\mathcal{E}^r_{eh}$-equivalence of real matrices. Thus, we have:

Corollary (6.7): Let $T_1, T_2 \in \mathbb{R}^{k\times 2}_n$, $k\geq 4$, and let $\tilde{\underline{g}}^1_{1,i_1,i_2}, \tilde{\underline{g}}^2_{1,i_1,i_2}$ be the $(1,i_1,i_2)$-$\mathbb{R}$-CGV of T_1, T_2 respectively. Necessary and sufficient condition for $T_1 \mathcal{E}^r_{eh} T_2$ is that

$$\tilde{\underline{g}}^2_{1,i_1,i_2} = \varepsilon \tilde{\underline{g}}^1_{1,i_1,i_2}, \ \varepsilon=\pm 1 \tag{6.51}$$

where (6.51) holds with "+","-", if and only if $T_1 \mathcal{E}^r_{eh} T_2$ with a $Q\in\mathbb{R}^{2\times 2}_n$ and such that $|Q|>0, |Q|<0$ correspondingly.

Proof

By following similar arguments as in the proof of Theorem (6.4), we obtain a modified form of eqn(6.50c) as

$$T^2_{1,i_1,i_2} = \frac{\varepsilon}{\sqrt{|\delta|}} T^1_{1,i_1,i_2} Q \tag{6.52a}$$

By applying the Binet-Cauchy theorem we have that

$$\tilde{\underline{g}}^{2}_{1,i_1,i_2} = \mathcal{C}_2\left(\frac{\varepsilon}{\sqrt{|\delta|}} T^{1}_{1,i_1,i_2} Q\right) = \frac{1}{|\delta|} \tilde{\underline{g}}^{1}_{1,i_1,i_2} |Q| = \varepsilon \tilde{\underline{g}}^{1}_{1,i_1,i_2} \qquad (6.25b)$$

Clearly, if $\delta>0$ then (6.51) holds with "+", and if $\delta<0$ it holds with "-". The sufficiency follows by a mere reversion of the arguments as in Theorem (6.4).

□

Remark (6.14): The $(1,i_1,i_2)$-$\mathbb{R}$-CGV for matrices of $\mathbb{R}^{k\times 2}_{n}$, $k\geq 4$ is a complete invariant for $\mathcal{E}^{r}_{eh}$-equivalence modulo ε, where $\varepsilon=+1,-1$ if and only if $\mathcal{E}^{r}_{eh}$-equivalence is defined with $|Q|>0, |Q|<0$ respectively transformations.

The role of the (r,i_1,i_2)-$\mathbb{R}$-CGV for $\mathcal{E}^{r}_{eh}$-equivalence defined on matrices of $\mathbb{C}^{k\times 3}_{n}$ is examined next. We may state the following:

Theorem (6.5): Let $T_1,T_2\in\mathbb{C}^{k\times 2}_{3}$, $k\geq 3$, and let $\tilde{\underline{g}}^{1}_{1,i_1,i_2}, \tilde{\underline{g}}^{2}_{1,i_1,i_2}$ be the $(1,i_1,i_2)$-$\mathbb{R}$-CGV of T_1,T_2 respectively. Necessary and sufficient condition for $T_1\mathcal{E}^{r}_{eh}T_2$ is that either of the following conditions hold true:

(i) $$\tilde{\underline{g}}^{2}_{1,i_1,i_2} = \tilde{\underline{g}}^{1}_{1,i_1,i_2} \qquad (6.53a)$$

or

(ii) $$\tilde{\underline{g}}^{2}_{1,i_1,i_2} = -\mathcal{C}_2(\bar{E})\tilde{\underline{g}}^{1}_{1,i_1,i_2} \qquad (6.53b)$$

where $\bar{E}=\text{diag}\{E,\dots,E\}$ and $E=[j]_{\mathbb{R}}$.

Proof

Assume $T_1\mathcal{E}^{r}_{eh}T_2$. By using similar arguments as in the proof of Theorem (6.4) we obtain eqn(6.50c), i.e.

$$T^{2}_{1,i_1,i_2} = \frac{\varepsilon}{\sqrt{\delta}} T^{1}_{1,i_1,i_2} Q, \quad |Q|=\delta \qquad (6.54a)$$

(i) If $T_1\mathcal{E}^{r}_{eh}T_2$ with $|Q|>0$, then (6.54a) implies that

$$F^{2}_{1,i_1,i_2} = \frac{\varepsilon}{\sqrt{\delta}} F^{1}_{1,i_1,i_2} Q \qquad (6.54b)$$

where $F^{i}_{1,i_1,i_2}=[T^{i}_{1,i_1,i_2}]_{\mathbb{R}}$, $i=1,2$. By the Binet-Cauchy Theorem we have that

$$\tilde{\underline{g}}^2_{1,i_1,i_2} = \mathcal{C}_2\left(\frac{\varepsilon}{\sqrt{\delta}} F^1_{1,i_1,i_2} Q\right) = \frac{1}{\delta}\tilde{\underline{g}}^1_{1,i_1,i_2}|Q| = \tilde{\underline{g}}^1_{1,i_1,i_2} \qquad (6.54c)$$

(ii) If $T_1 \mathcal{E}^r_{eh} T_2$ with $|Q|<0$, then $\sqrt{\delta}=j\sqrt{-\delta}$ and $1/\sqrt{\delta}=-j/\sqrt{-\delta}$. Thus, (6.54b) implies

$$F^2_{1,i_1,i_2} = \frac{\varepsilon}{\sqrt{-\delta}} \bar{E}\, F^1_{1,i_1,i_2} Q \qquad (6.54d)$$

where $\bar{E}=\text{diag}\{E,\dots,E\}$, $E=[j]_{\mathbb{R}}$. By the Binet-Cauchy Theorem, it follows that

$$\tilde{\underline{g}}^2_{1,i_1,i_2} = \mathcal{C}_2\left(\frac{\varepsilon}{\sqrt{-\delta}} \bar{E}\, F^1_{1,i_1,i_2} Q\right) = \frac{1}{(-\delta)}\,\mathcal{C}_2(\bar{E})\tilde{\underline{g}}^1_{1,i_1,i_2}|Q| = -\mathcal{C}_2(\bar{E})\tilde{\underline{g}}^1_{1,i_1,i_2} \qquad (6.54e)$$

To prove sufficiency assume that (6.54c) or (6.54e) hold true. Then, by the decomposability of $\tilde{\underline{g}}^2_{1,i_1,i_2}, \tilde{\underline{g}}^1_{1,i_1,i_2}$ and $\mathcal{C}_2(\bar{E})$ we have that

$$(6.54c) \iff F^2_{1,i_1,i_2} = F^1_{1,i_1,i_2} Q' \text{ with } Q' \in \mathbb{R}^{2\times 2}_n \text{ and } |Q'|=1 \qquad (6.54f)$$

$$(6.54e) \iff F^2_{1,i_1,i_2} = \bar{E}F^1_{1,i_1,i_2} Q' \text{ with } Q' \in \mathbb{R}^{2\times 2}_n \text{ and } |Q'|=-1 \qquad (6.54g)$$

However, by translating (6.54f) and (6.54g) into the complex form we have:

$$(6.54f) \iff T^2_{1,i_1,i_2} = T^1_{1,i_1,i_2} Q',\ |Q'|=1 \qquad (6.54h)$$

$$(6.54g) \iff T^2_{1,i_1,i_2} = jT^1_{1,i_1,i_2} Q',\ |Q'|=-1 \qquad (6.54k)$$

Both of the above cases imply that $T^2_{1,i_1,i_2} \mathcal{E}^r_{eh} T^1_{1,i_1,i_2}$ and thus $T_1 \mathcal{E}^r_{eh} T_2$. □

Corollary (6.8): The $(1,i_1,i_2)$-$\mathbb{R}$-CGV, $\tilde{\underline{g}}_{1,i_1,i_2}$, of $T\in\mathbb{C}^{k\times 2}_n$, or any other (r,j_1,j_2)-$\mathbb{R}$-CGV $\tilde{\underline{g}}_{r,j_1,j_2}$, is a complete invariant modulo $I_{\binom{2k}{2}}$, or modulo $-\mathcal{C}_2(\bar{E})$, for $\mathcal{E}^r_{eh}$-equivalence. In particular:

(i) $\tilde{\underline{g}}^2_{1,i_1,i_2} = \tilde{\underline{g}}^1_{1,i_1,i_2}$ for $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ with $|Q|>0$

(ii) $\tilde{\underline{g}}^2_{1,i_1,i_2} = -\mathcal{C}_2(\bar{E})\tilde{\underline{g}}^1_{1,i_1,i_2}$ for $\beta\in PGL(1,\mathbb{C}/\mathbb{R})$ with $|Q|<0$

Remark (6.15): The set of r-prime Plücker vectors $\mathcal{P}_r$, or equivalently the (r,i_1,i_2)-$\mathbb{C}$-CGV, $\underline{g}_{r,i_1,i_2}$, are dependent invariants on the (r,i_1,i_2)-$\mathbb{R}$-CGV

$\tilde{g}_{r,i_1,i_2}$ for $\mathcal{E}^r_{eh}$-equivalence. Thus, for the characterisation of $\mathcal{E}^r_{eh}$-equivalence we need only one complete invariant, the $\tilde{g}_{r,i_1,i_2}$ vector, for some $r\in<k>$ and $(i_1,i_2)\in Q^{r)}_{2,k-1}$ fixed.

Note that the vectors g_{r,i_1,i_2} and $\tilde{g}_{r,i_1,i_2}$ uniquely characterise families of matrices $\{T_{r,i_1,i_2}\}$ and $\{F_{r,i_1,i_2}\},\{\bar{E}F_{r,i_1,i_2}\}$ with the properties:

$$T'_{r,i_1,i_2} = T_{r,i_1,i_2}Q, \ |Q|=1, \quad \forall\, T'_{r,i_1,i_2}\in\{T_{r,i_1,i_2}\} \tag{6.55}$$

$$\begin{cases} F_{r,i_1,i_2} = F_{r,i_1,i_2}Q, \ |Q|=1, & \forall\, F'_{r,i_1,i_2}\in\{F_{r,i_1,i_2}\} \\ F_{r,i_1,i_2} = \bar{E}F_{r,i_1,i_2}Q, \ |Q|=-1, & \forall\, F'_{r,i_1,i_2}\in\{\bar{E}F_{r,i_1,i_2}\} \end{cases} \tag{6.56}$$

The vector spaces $\mathrm{span}_{\mathbb{C}}\{T_{r,i_1,i_2}\}\triangleq \mathcal{T}_{r,i_1,i_2}$ and $\mathrm{span}_{\mathbb{R}}\{F_{r,i_1,i_2}\}\triangleq \mathcal{F}_{r,i_1,i_2}$, $\mathrm{span}_{\mathbb{R}}\{\bar{E}F_{r,i_1,i_2}\}\triangleq \mathcal{F}^*_{r,i_1,i_2}$ are invariant under $\mathcal{E}^c_{eh}$-,$\mathcal{E}^r_{eh}$-equivalence and are defined as the <u>$\mathcal{E}^c_{eh}$-, and $\mathcal{E}^r_{eh}$-equivalence characteristic spaces</u> of T respectively. The interpretation of the $g_{r,i_1,i_2},\tilde{g}_{r,i_1,i_2}$ on the corresponding (r,i_1,i_2)-$\mathbb{C}$, -$\mathbb{R}$-standard forms leads to the following result, that concludes this section.

<u>Corollary (6.9)</u>: Let $T_1,T_2\in\mathbb{C}^{k\times 2}_n$, k≥3, and let $\{T^1_{r,i_1,i_2},F^1_{r,i_1,i_2}\}$, $\{T^1_{r,i_1,i_2},F^2_{r,i_1,i_2}\}$ be the (r,i_1,i_2)-$\mathbb{C}$, $\mathbb{R}$-standard forms of T_1,T_2 respectively. Then

(i) $T_1\mathcal{E}^c_{eh}T_2$, iff $T^1_{r,i_1,i_2}\,\mathcal{E}^c_r\,T^2_{r,i_1,i_2}$ with $|Q|=1$ (right complex equivalent).

(ii) $T_1\mathcal{E}^r_{eh}T_2$, iff $F^1_{r,i_1,i_2}\,\mathcal{E}^r_r\,F^2_{r,i_1,i_2}$ with $|Q|=1$, or $\bar{E}F^1_{r,i_1,i_2}\,\mathcal{E}^r_r\,F^2_{r,i_1,i_2}$ with $|Q|=-1$ (right real equivalent).

□

6.6 A complete set of invariants of homogeneous binary polynomials under projective equivalence

The necessary and sufficient conditions for $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})\in\mathbb{R}_d\{\Theta\}$ to be $\mathcal{E}_P$ equivalent, have been given by Theorem (6.2); in this result the notion of $\mathcal{E}^r_{eh}$-equivalence is of crucial importance. The characterisation of the

$\mathcal{E}^{r}_{eh}(T)$, $T\in\mathbb{C}_{n}^{k\times 2}$, by a complete set of invariants provides the basis for defining a complete set of invariants for $\mathcal{E}_{\mathcal{P}}(f)$ and thus for $\mathcal{E}_{\mathcal{P}}(F)$. Theorem (6.1) may then be used for the derivation of a complete set of invariants for the $\mathcal{E}_{\mathcal{H}\text{-}\mathcal{B}}(F,G)$ equivalence class. In this section, the cases of $\mathcal{E}_{\mathcal{P}}(f)$ and $\mathcal{E}_{\mathcal{P}}(F)$ equivalence classes are considered here. The original problem, the study of invariants of $\mathcal{E}_{\mathcal{H}\text{-}\mathcal{B}}(F,G)$ is examined in the next section. We first define the following:

Definition (6.6): Let $f(s,\hat{s})\in\mathbb{R}_{k}\{\theta\}$, $\mathcal{B}(f)=\{\mathcal{B}_{\mathbb{R}}(f);\mathcal{B}_{\mathbb{C}}(f)\}$ and let $\mu\triangleq\#\mathcal{B}_{\mathbb{R}}(f)$, $\nu\triangleq\#\mathcal{B}_{\mathbb{C}}(f)$ (# denotes the number of elements of the corresponding set). Let us also assume that $\pi\in<\mathcal{B}_{\mathbb{R}}(f)>$, $\pi'\in<\mathcal{B}_{\mathbb{C}}(f)>$ and that

$$T_{\pi,\pi'} = [\mathcal{B}^{\pi,\pi'}(f)] = \begin{bmatrix} [\mathcal{B}^{\pi}_{\mathbb{R}}(f)] \\ [\mathcal{B}^{\pi'}_{\mathbb{C}}(f)] \end{bmatrix} \qquad (6.57)$$

is the (π,π')-basis matrix of $f(s,\hat{s})$. We may define:

(i) $\mathcal{T}(f)\triangleq\{T_{\pi,\pi'}:\forall\,\pi\in<\mathcal{B}_{\mathbb{R}}(f)>,\ \forall\,\pi'\in<\mathcal{B}_{\mathbb{C}}(f)>\}$ as the family of matrix representations of $\mathcal{B}(f)$, or of $f(s,\hat{s})$. The ordered pair (μ,ν) will be referred to as the order of $\mathcal{B}(f)$, or of $f(s,\hat{s})$.

(ii) If $\mu+\nu\geq 3$, $T_{\pi,\pi'}\in\mathcal{T}(f)$, $r\in<\mu+\nu>$, $(i_1,i_2)\in Q^{r)}_{2,\mu+\nu-1}$, then the (r,i_1,i_2)-$\mathbb{R}$-CGV of $T_{\pi,\pi'}$, $\tilde{\underline{g}}^{\pi,\pi'}_{r,i_1,i_2}$, is well defined and shall be referred to as the (π,π')-(r,i_1,i_2)-$\mathbb{R}$-canonical-Grassmann vector ((π,π')-(r,i_1,i_2)-$\mathbb{R}$-CGV) of $f(s,\hat{s})$. The set of all such vectors

$$G_{f}\triangleq\{\tilde{\underline{g}}^{\pi,\pi'}_{r,i_1,i_2}:\forall\,\pi\in<\mathcal{B}_{\mathbb{R}}(f)>,\ \forall\,\pi'\in<\mathcal{B}_{\mathbb{C}}(f)>,\ \forall\,r\in<\mu+\nu> \text{ and } \forall\,(i_1,i_2)\in Q^{r)}_{2,\mu+\nu-1}\} \qquad (6.58)$$

is well defined and shall be called the $\mathbb{R}$-canonical Grassmann vector set ($\mathbb{R}$-CGVS) of $f(s,\hat{s})$. □

Note that the set G_f is completely defined by an element $\underline{g}^{\pi,\pi'}_{r,i_1,i_2}$. In fact, $\underline{g}^{\pi,\pi'}_{r,i_1,i_2}$ defines a family of (r,i_1,i_2)-standardised matrices

$\{T_{\pi,\pi'}^{r,i_1,i_2}\}$ which are right equivalent with transformations Q, having $|Q|=1$. From any $T\in\{T_{\pi,\pi'}^{r,i_1,i_2}\}$, we may construct all other elements of G_f; this is why any $g_{r,i_1,i_2}^{\pi,\pi'}$ may be referred to as a <u>generator</u> of G_f.

With the above definitions in mind, we may now state the conditions for $\mathcal{E}_p$-equivalence of elements of $\mathbb{R}_d\{\Theta\}$. The simple cases are examined first.

<u>Proposition (6.16)</u>: Let $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$, $\mathcal{B}(f)=\{\mathcal{B}_{\mathbb{R}}(f),\mathcal{B}_{\mathbb{C}}(f)\}$, $J_{\mathbb{R}}(f),J_{\mathbb{C}}(f)$ be the sets associated with $f(s,\hat{s})$ and let (μ,ν) be the order of $f(s,\hat{s})$. The sets $J_{\mathbb{R}}(f),J_{\mathbb{C}}(f)$ and the order (μ,ν) form a complete set of invariants for $\mathcal{E}_p(f)$ in the following cases:

(i) $(\mu,\nu)=(1,0)$, $(\mu,\nu)=(2,0)$, $(\mu,\nu)=(3,0)$

(ii) $(\mu,\nu)=(0,1)$, $(\mu,\nu)=(1,1)$ □

The proof of the above result readily follows by Theorem (6.2) and Proposition (6.14). For polynomials with $(\mu,\nu)=(0,2)$, $\mathcal{B}(f)=\mathcal{B}_{\mathbb{C}}(f)$ and $J(f)=J_{\mathbb{C}}(f)$; thus, if $\mathcal{B}_{\mathbb{C}}(f)=\{(\gamma_1,\delta_1;p_1),(\gamma_2,\delta_2;p_2),p_1\le p_2\}$, $J_{\mathbb{C}}(f)=\{(p_1,1),(p_2,1)\}$ and if we denote by $e_i=(\gamma_i,\delta_i)$, $[e_i]_{\mathbb{R}}=\Sigma_i$, $i=1,2$ and by

$$A=\Sigma_2\Sigma_1^{-1}=\begin{bmatrix}a & b\\ c & d\end{bmatrix},\ \bar{A}=\Sigma_1\Sigma_2^{-1}=\begin{bmatrix}\bar{a} & \bar{b}\\ \bar{c} & \bar{d}\end{bmatrix} \tag{6.59}$$

then we may define the following characteristic matrices of f:

$$P^{\alpha}=\begin{bmatrix}a & b\\ b & -a\\ c & d\\ d & -c\end{bmatrix},\ P^{\beta}=\begin{bmatrix}a & c\\ b & d\\ c & -a\\ d & -b\end{bmatrix},\ P^{\bar{\alpha}}=\begin{bmatrix}\bar{a} & \bar{b}\\ \bar{b} & -\bar{a}\\ \bar{c} & \bar{d}\\ \bar{d} & -\bar{c}\end{bmatrix},\ P^{\bar{\beta}}=\begin{bmatrix}\bar{a} & \bar{c}\\ \bar{b} & \bar{d}\\ \bar{c} & -\bar{a}\\ \bar{d} & -\bar{b}\end{bmatrix} \tag{6.60}$$

The above matrices will be referred to as the <u>α-,β-,$\bar{\alpha}$-,$\bar{\beta}$-characteristic matrices</u> of $f(s,\hat{s})$ and they are matrices of $\mathbb{R}^{4\times 2}$. If $T_1,T_2\in\mathbb{R}^{4\times 2}$, then the property described by $|[T_1,T_2]|=0$ is denoted by $T_1\square T_2=0$. With these definitions in mind, we may state the following result:

Proposition (6.17): Let $f_1, f_2 \in \mathbb{R}_d\{\Theta\}$, $(\mu,\nu)=(0,2)$ and let $J_{\mathbb{C}}(f_i)=\{(p_1^i,1), (p_2^i,1)\}$ and $P_i^{\alpha}, P_i^{\beta}, P_i^{\bar{\alpha}}, P_i^{\bar{\beta}}$, $i=1,2$, be the corresponding complex list and characteristic matrices of f_i. Necessary and sufficient conditions for $f_i \mathcal{E}_{\mathcal{P}} f_2$ are:

(i) $J_{\mathbb{C}}(f_1) = J_{\mathbb{C}}(f_2)$.

(ii) (α) If $p_1^1 < p_2^1$, then $P_1^{\beta} \square P_2^{\alpha} = 0$.

(β) If $p_1^1 = p_2^1$, then either $P_1^{\beta} \square P_2^{\alpha} = 0$, or $P_1^{\bar{\beta}} \square P_2^{\bar{\alpha}} = 0$.

□

The proof readily follows by Theorem (6.2) and Proposition (6.15). Our attention is focussed next on the general case, which is not covered by Propositions (6.16) and (6.17). A polynomial of $\mathbb{R}_d\{\Theta\}$ for which the order (μ,ν) is different from $(1,0),(2,0),(3,0),(0,1),(1,1)$ and $(0,2)$ will be referred to as a general order polynomial. For such elements of $\mathbb{R}_d\{\Theta\}$ we have the following result:

Theorem (6.6): Let $f(s,\hat{s}) \in \mathbb{R}_d\{\Theta\}$ be a general order polynomial and let $\mathcal{B}(f)=\{\mathcal{B}_{\mathbb{R}}(f), \mathcal{B}_{\mathbb{C}}(f)\}$, $J(f)=\{J_{\mathbb{R}}(f), J_{\mathbb{C}}(f)\}$ and $\mathcal{G}_f$ be the sets associated with $f(s,\hat{s})$. A complete set of invariants for the $\mathcal{E}_{\mathcal{P}}(f)$-equivalence class is defined by:

(i) $J_{\mathbb{R}}(f), J_{\mathbb{C}}(f)$.

(ii) $\mathcal{G}_f$, or $\mathcal{G}_f$ modulo $-\mathcal{C}_2(\bar{E})$, where $\bar{E}=\text{diag}\{E,\ldots,E\}$ and $E=[j]_{\mathbb{R}}$.

Proof

If $f_1, f_2 \in \mathbb{R}_d\{\Theta\}$ are two general order polynomials and $f_1 \mathcal{E}_{\mathcal{P}} f_2$, then by Theorem (6.2) we have that $(\mu_1,\nu_1)=(\mu_2,\nu_2)$, $J_{\mathbb{R}}(f_1)=J_{\mathbb{R}}(f_2)$ $J_{\mathbb{C}}(f_1)=J_{\mathbb{C}}(f_2)$ and there exist $\pi_i(\mathcal{B}_{\mathbb{R}}(f_i))$, $\pi_i'(\mathcal{B}_{\mathbb{C}}(f_i)$, $i=1,2$, such that $T^2_{\pi_2,\pi_2'}=\text{diag}\{\xi_i\}T^1_{\pi_1,\pi_1'}Q$, where $\xi_i \in \mathbb{C}-\{0\}$ and $Q \in \mathbb{R}_n^{2\times 2}$.

By the last condition ($\mathcal{E}^r_{eh}$-equivalence of $T^1_{\pi_1,\pi_1'}, T^2_{\pi_2,\pi_2'}$) and by Theorem (6.5) it follows that

$$\tilde{g}^{\pi_2,\pi_2'}_{r,i_1,i_2} = \tilde{g}^{\pi_1,\pi_1'}_{r,i_1,i_2} \text{ or } -\mathcal{C}_2(\bar{E})\tilde{g}^{\pi_1,\pi_1'}_{r,i_1,i_2} \tag{6.61}$$

Note that eqn(6.61) holds true for all (r,i_1,i_2); furthermore, $T^1_{\pi_1,\pi_1'} E^r_{eh} T^2_{\pi_2,\pi_2'}$ implies that for ∀ $(\tilde{\pi}_1,\tilde{\pi}_1')$ we may find a suitable $(\tilde{\pi}_2,\tilde{\pi}_2')$ for which $T^1_{\tilde{\pi}_1,\tilde{\pi}_1'} E^r_{eh} T^2_{\tilde{\pi}_2,\tilde{\pi}_2'}$; thus, by (6.61), $G_{f_1}=G_{f_2}$ or $G_{f_1}=G_{f_2}$ (mod-$C_2(\bar{E})$). This proves the necessity (invariance). To prove the sufficiency (completeness) assume $J_{\mathbb{R}}(f_1)=J_{\mathbb{R}}(f_2)$, $J_{\mathbb{C}}(f_1)=J_{\mathbb{C}}(f_2)$ and that $G_{f_1}=G_{f_2}$, or $G_{f_1}=G_{f_2}$ (mod-$C_2(\bar{E})$). The last condition implies that there exists a one to one correspondence between the elements of G_{f_1}, G_{f_1} which is expressed by eqn(6.61). Choose an (r,i_1,i_2), and pairs $(\pi_1,\pi_1'),(\pi_2,\pi_2')$ for which the (6.61) type relations hold true. Then by Theorem (6.5) we have that $T^1_{\pi_1,\pi_1'} E^r_{eh} T^2_{\pi_2,\pi_2'}$ and by Theorem (6.2) the completeness is established. □

Remark (6.16): If G_{f_1},G_{f_2} are two ℝ-CGVS, $\underline{g}_1 \in G_{f_1}$, $\underline{g}_2 \in G_{f_2}$ and either $\underline{g}_2=\underline{g}_1$, or $\underline{g}_2=-C_2(\bar{E})\underline{g}_1$, then $G_{f_2}=G_{f_1}$ or $G_{f_2}=G_{f_1}$ (mod-$C_2(\bar{E})$) respectively. Thus, two sets G_{f_1},G_{f_2} are equal, or mod($-C_2(\bar{E})$) equal, if and only if they have a common, or mod($-C_2(\bar{E})$) common point correspondingly. □

In a manner similar to that described for the construction of G_f, we may construct the set of all (π,π')-(r,i_1,i_2)-ℂ-CGVS, $\underline{g}^{\pi,\pi'}_{r,i_1,i_2}$, which shall be denoted by $\bar{G}_f$. For every $T_{\pi,\pi'} \in T(f)$, we may also construct the set of all r-prime Plücker vectors, $P^{\pi,\pi'}$, by $P^{\pi,\pi'} = \bigcup_{i=1}^{k} P_i^{\pi,\pi'}$, where $P_i^{\pi,\pi'}$ is the i-prime Plucker vector set, where $k=\mu+\nu$. Then, the set

$$P^* = \bigcup_{\pi,\pi'} P^{\pi,\pi'}, \text{ for } \forall\ \pi \in \langle B_{\mathbb{R}}(f)\rangle,\ \forall\ \pi' \in \langle B_{\mathbb{C}}(f)\rangle \qquad (6.62)$$

will be called the Plücker-vector set of $f(s\hat{,}s)$. For polynomials with real roots we have the following result:

Corollary (6.10): Let $f(s,\hat{s}) \in \mathbb{R}_d\{\Theta\}$ be a general order polynomial with real roots and let $B(f)=B_{\mathbb{R}}(f)$, $J(f)=J_{\mathbb{R}}(f)$, G_f and P^*_f be the sets associated with $f(s,\hat{s})$. A complete set of invariants for $E_P(f)$-equivalence class is defined by:

(i) $J_{\mathbb{R}}(f)$.

(ii) G_f modulo ±1, or equivalently $\mathcal{P}_f^*$ modulo collinearity of the corresponding vectors.

The proof of the above result follows immediately by Theorem (6.2), Corollary (6.7) and from the equivalence of the set of r-prime Plücker vectors $\mathcal{P}_r$, to the (r,i_1,i_2)-$\mathbb{R}$-CGV g_{r,i_1,i_2}, established by Corollary (6.6).

Remark (6.17): If $\mathcal{P}_1^*,\mathcal{P}_2^*$ are two sets of Plücker vector sets, $\mathcal{P}_1^{\pi,\pi'}\in\mathcal{P}_1^*$, $\mathcal{P}_2^{\tilde{\pi},\tilde{\pi}'}\in\mathcal{P}_2^*$ and $\mathcal{P}_1^{\pi,\pi'},\mathcal{P}_2^{\tilde{\pi},\tilde{\pi}'}$ collinear, then $\mathcal{P}_1^*,\mathcal{P}_2^*$ are collinear. Thus, $\mathcal{P}_1^*,\mathcal{P}_2^*$ are equal modulo collinearity of the corresponding vectors, if and only if they have a common (modulo collinearity of the corresponding vectors) vector set $\mathcal{P}_1^{\pi,\pi'}=\mathcal{P}_2^{\tilde{\pi},\tilde{\pi}'}=\mathcal{P}^{\pi,\pi'}$.

We may illustrate the notion of $\mathcal{E}_{\mathcal{P}}$-equivalence on $\mathbb{R}_d\{\Theta\}$ by the following example.

Example (6.3): Let $f(s,\hat{s}),\tilde{f}(\lambda,\hat{\lambda})\in\mathbb{R}_5\{\Theta\}$, where

$$f(s,\hat{s}) = (s-3\hat{s})^2(s-2\hat{s})(2s-3\hat{s})(s-4\hat{s})$$

$$\tilde{f}(\lambda,\hat{\lambda}) = (4\lambda-11\hat{\lambda})^2(6\lambda-16\hat{\lambda})(15\lambda-39\hat{\lambda})(20\lambda-56\hat{\lambda})$$

We shall examine whether $f(s,\hat{s})\mathcal{E}_{\mathcal{P}}\tilde{f}(\lambda,\hat{\lambda})$. Given that both polynomials have real roots, the UFS of $f(s,\hat{s})$ and $\tilde{f}(\lambda,\hat{\lambda})$ are given by

$$\begin{aligned} \mathcal{B}(f) &= \mathcal{B}_{\mathbb{R}}(f) = \{(1,2;1),(2,3;1)(1,4;1);(1,3;2)\} \\ \mathcal{B}(\tilde{f}) &= \mathcal{B}_{\mathbb{R}}(\tilde{f}) = \{(6,16;1),(15,39;1),(20,56;1),(4,11;2)\} \end{aligned} \tag{6.63a}$$

and thus

$$J_{\mathbb{R}}(f) = \{(3,1),(1,2)\} = J_{\mathbb{R}}(\tilde{f}) \tag{6.63b}$$

Since the first of the two conditions of Corollary (6.10) is satisfied, we proceed to the checking of the second. Select the following permutations from $\mathcal{B}(f)$ and $\mathcal{B}(\tilde{f})$:

$$\begin{aligned}\pi(\mathcal{B}(f)) &= \{(1,2;1),(1,4;1),(2,3;1);(1,3;2)\}\\ \pi'(\mathcal{B}(\tilde{f})) &= \{(6,16;1),(15,39;1),(20,56;1);(4,11;2)\}\end{aligned} \tag{6.63c}$$

Then,

$$T_1^{\pi} = [\pi(\mathcal{B}(f))] = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 2 & 3 \\ 1 & 3 \end{bmatrix}, \; T_2^{\pi'} = [\pi'(\mathcal{B}(f))] = \begin{bmatrix} 6 & 16 \\ 15 & 39 \\ 20 & 56 \\ 4 & 11 \end{bmatrix} \tag{6.63d}$$

Let $r=1\in<4>$. Then there exists one quadruple $\phi=(1,2,3,4)$ based on 1 and so one prime triangle $\delta\phi_1=\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$. The corresponding Grassmann vectors for $T_1^{\pi}, T_2^{\pi'}$ are:

$$\underline{g}(T_1^{\pi}) = [2,-1,1,-5,-1,3]^t, \; \underline{g}(T_2^{\pi'}) = [-6,16,2,60,9,-4]^t \tag{6.63e}$$

and thus the Plücker vectors for r=1 are

$$\underline{p}_1^{\pi}(\delta\phi_1) = \begin{bmatrix} 6 \\ -1 \\ -5 \end{bmatrix}, \; \underline{p}_2^{\pi'}(\delta\phi_1) = \begin{bmatrix} 24 \\ -144 \\ 120 \end{bmatrix} = -24 \begin{bmatrix} -1 \\ 6 \\ 5 \end{bmatrix} \tag{6.63f}$$

Thus, since $\underline{p}_1^{\pi}(\delta\phi_1)\neq\lambda\underline{p}_2^{\pi'}(\delta\phi_1)$, the matrices $T_1^{\pi}, T_2^{\pi'}$ are not $\mathcal{E}_{eh}^{r}$-equivalent. Eqn(6.63f), however, suggests that there may exist a different permutation on $\mathcal{B}(\tilde{f})$ for which $\mathcal{E}_{eh}^{r}$ may be established. Thus, let us take

$$\pi''(\mathcal{B}(\tilde{f})) = \{(6,15;1),(20,56;1),(15,39-1);(4,11;2)\} \tag{6.63g}$$

Then,

$$T_2^{\pi''} = [\pi''(\mathcal{B}(\tilde{f}))] = \left[\begin{array}{cc} 6 & 16 \\ 20 & 56 \\ 15 & 39 \\ \hline 4 & 11 \end{array}\right] \text{ and } \underline{g}(T_2^{\pi''}) = [16,-6,2,-60,-4,9]^t \tag{6.63h}$$

and thus

$$\underline{p}_2^{\pi''}(\delta\phi_1) = \begin{bmatrix} 144 \\ -24 \\ -120 \end{bmatrix} = 24\underline{p}_1^{\pi}(\delta\phi_1) \qquad (6.63k)$$

The above condition establishes the collinearity for the ordered pair (π,π'') of permutations; thus $T_1^{\pi} E_{eh}^{r} T_2^{\pi''}$ and $f(s,\hat{s}) E_{\mathcal{P}} \tilde{f}(\lambda,\hat{\lambda})$. □

The characterisation of the $E_{\mathcal{P}}(f)$-equivalence class by a complete set of invariants, provides also a solution to the general linear mapping problem. In fact, the UFS $\mathcal{B}(f)$ of an $f(s) \in \mathbb{R}_d\{\Theta\}$ defines a general symmetric set points ζ of $\mathbb{C}\cup\{\infty\}$, where the list $J(f)$ defines the corresponding multiplicities of the distinct points of ζ. Because of this observation we have:

Remark (6.12): The general linear mapping problem is equivalent to the study of complete invariants of polynomials of $\mathbb{R}_d\{\Theta\}$ under $E_{\mathcal{P}}$-equivalence. The necessary and sufficient conditions for $f_1 E_{\mathcal{P}} f_2$, $f_1, f_2 \in \mathbb{R}_d\{\Theta\}$, are also necessary and sufficient conditions for the solution of the GLMP, since every symmetric point set ζ may be considered as the UFS of some $f \in \mathbb{R}_d\{\Theta\}$. □

The results obtained so far for $E_{\mathcal{P}}$-equivalence of an $f(s,\hat{s}) \in \mathbb{R}_d\{\Theta\}$ are extended next to the case of a set of polynomials. Of special interest are sets of binary polynomials ordered by the divisibility property. The study of $E_{\mathcal{P}}$-equivalence on such sets allows the connection of the results of Section (6.3) with those of (6.5) and thus yield a complete set of invariants of matrix pencils under $E_{\mathcal{H}\text{-}\mathcal{B}}$-equivalence.

Let $\mathcal{J}=\{d_i : i\in\underset{\sim}{\rho}, d_1 \le d_2 \le \ldots \le d_\rho\}$, $\mathbb{R}_{\mathcal{J}}\{\Theta\} = \mathbb{R}_{d_1}\{\Theta\}\times\ldots\times\mathbb{R}_{d_\rho}\{\Theta\}$, and let $\mathcal{F}=(f_1(s,\hat{s}),\ldots,f_\rho(s,\hat{s})) \in \mathbb{R}_{\mathcal{J}}\{\Theta\}$; $\mathcal{F}$ will be said to be a Smith-type ρ-tuple, if and only if it is ordered by divisibility, i.e. $f_i(s,\hat{s})/f_{i+1}(s,\hat{s})$, for $\forall\ i\in\underset{\sim}{\rho-1}$. The set of all Smith-type ρ-tuples $\mathcal{F}$ of $\mathbb{R}_{\mathcal{J}}\{\Theta\}$ will be called a $\mathcal{J}$-Smith family and shall be denoted by $S\mathbb{R}_{\mathcal{J}}\{\Theta\}$. For every $\mathcal{F} \in S\mathbb{R}_{\mathcal{J}}\{\Theta\}$ we can define the set of elementary divisors of $\mathcal{F}$ by factorising over $\mathbb{C}$ ($\mathbb{R}$) every $f_i(s,\hat{s})$ of $\mathcal{F}$; the set of e.d. of $\mathcal{F}$ over $\mathbb{C}$ will be denoted by $\mathcal{D}_{\mathcal{F}}$, and it is clearly a symmetric set (e.d. appear in complex conjugate pairs).

Lemma (6.3): [Gant - 1] Let $\mathcal{F}\epsilon S\mathbb{R}_{\mathcal{J}}\{\Theta\}$ and let $\mathcal{D}_{\mathcal{F}}$ be the set of e.d. of $\mathcal{F}$ over $\mathbb{C}$. The set $\mathcal{D}_{\mathcal{F}}$ uniquely defines the elements of modulo scaling by nonzero constants. □

This standard result provides a representation for $\mathcal{F}$ by the set $\mathcal{D}_{\mathcal{F}}$. The following result emphasises the importance of the set $\mathcal{D}_{\mathcal{F}}$.

Proposition (6.18): Let $\mathcal{F}_1,\mathcal{F}_2\epsilon S\mathbb{R}_{\mathcal{J}}\{\Theta\}$ and let $\mathcal{D}_{\mathcal{F}_1},\mathcal{D}_{\mathcal{F}_2}$ be the corresponding e.d. sets. $\mathcal{F}_1 \mathcal{E}_{\mathcal{P}} \mathcal{F}_2$ if and only if there exists an ordering of the elements of $\mathcal{D}_{\mathcal{F}_1},\mathcal{D}_{\mathcal{F}_2}$, denoted by $\hat{\mathcal{D}}_{\mathcal{F}_1},\hat{\mathcal{D}}_{\mathcal{F}_2}$, such that for this ordering $\hat{\mathcal{D}}_{\mathcal{F}_1} \mathcal{E}_{\mathcal{P}} \hat{\mathcal{D}}_{\mathcal{F}_2}$. □

The proof of this proposition readily follows by Lemma (6.3) and Proposition (6.6). Thus, $\mathcal{E}_{\mathcal{P}}$-equivalence on Smith-type tuples $\mathcal{F}$ may be reduced to an equivalent problem of $\mathcal{E}_{\mathcal{P}}$-equivalence on sets of e.d. over $\mathbb{C}$. The notion of the ordering of $\mathcal{D}_{\mathcal{F}}$ sets is important for our study and shall be discussed next. Note that if $\mathcal{F}$ is not a Smith-type tuple, then $\mathcal{D}_{\mathcal{F}}$ does not define uniquely the set $\mathcal{F}$.

The symmetric set $\mathcal{D}_{\mathcal{F}}$ may be written as $\mathcal{D}_{\mathcal{F}}=\{\mathcal{D}_{\mathcal{F}}^{\mathbb{R}};\mathcal{D}_{\mathcal{F}}^{\mathbb{C}}\}$, where $\mathcal{D}_{\mathcal{F}}^{\mathbb{R}},\mathcal{D}_{\mathcal{F}}^{\mathbb{C}}$ denote respectively the sets of real, complex e.d. in $\mathcal{D}_{\mathcal{F}}$. The sets $\mathcal{D}_{\mathcal{F}}^{\mathbb{R}},\mathcal{D}_{\mathcal{F}}^{\mathbb{C}}$ and thus $\mathcal{D}_{\mathcal{F}}$ may be ordered as it is explained next. Let us denote by $\mathcal{I}_{\tau}\triangleq\{d_i^{\tau}:d_i^{\tau}\epsilon\mathbb{Z},i\epsilon\underset{\sim}{\sigma}_{\tau},d_1^{\tau}\leq\ldots\leq d_{\sigma_{\tau}}^{\tau}\}$ be an ordered set of integers and let

$$\mathcal{D}^{\mathbb{C}}(\mathcal{I}_{\tau'}) \triangleq \{(\gamma_j^{\tau'}s-\delta_j^{\tau'}\hat{s})^k,(\bar{\gamma}_j^{\tau'}s-\bar{\delta}_j^{\tau'}\hat{s})^k,k\epsilon\mathcal{I}_{\tau'};j\epsilon\underset{\sim}{\nu}_{\tau'},\gamma_j^{\tau'},\delta_j^{\tau'}\epsilon\mathbb{C},$$

$$(\gamma_j^{\tau'},\delta_j^{\tau'})\neq\xi(\gamma_i^{\tau'},\delta_i^{\tau'}),\forall i\neq j,\xi\epsilon\mathbb{C}-\{0\}\} \tag{6.64}$$

$$\mathcal{D}^{\mathbb{R}}(\mathcal{I}_{\tau}) \triangleq \{(\alpha_j^{\tau}s-\beta_j^{\tau}\hat{s})^k,k\epsilon\mathcal{I}_{\tau};j\epsilon\underset{\sim}{\nu}_{\tau},\alpha_j^{\tau},\beta_j^{\tau}\epsilon\mathbb{R}$$

$$(\alpha_j^{\tau},\beta_j^{\tau})\ \xi(\alpha_i^{\tau},\beta_i^{\tau}),\forall i\neq j,\xi\,\epsilon\mathbb{R}-\{0\}\} \tag{6.65}$$

be the sets of all pairs of complex conjugate e.d., real e.d., which have the same set of degrees $I_{\tau'}, I_\tau$ respectively (or the same Segré characteristics $I_{\tau'}, I_\tau$ respectively). The index sets $I_\tau (I_{\tau'})$ may be ordered as follows: For two sets I_α, I_β with $\sigma_\alpha \neq \sigma_\beta$, we say that I_α strictly precedes I_β, written $I_\alpha \dot{<} I_\beta$, if $\sigma_\alpha < \sigma_\beta$; if $\sigma_\alpha = \sigma_\beta$, then we say that I_α precedes I_β, written $I_\alpha < I_\beta$, if there exists an integer t ($1 \le t \le \sigma_\alpha = \sigma_\beta$), for which $d_1^\alpha = d_1^\beta, \ldots, d_{t-1}^\alpha = d_{t-1}^\beta, d_t^\alpha < d_t^\beta$. Thus, $(3,4) \dot{<} (1,2,3)$ and $(1,3,4) < (1,3,5)$. This notion of ordering for sets of integers, implies an ordering for $\mathcal{D}_F^{\mathbb{C}}, \mathcal{D}_F^{\mathbb{R}}$, and thus for $\mathcal{D}_F$. Thus, let $\mathcal{D}_F^{\mathbb{K}}$ be either $\mathcal{D}_F^{\mathbb{R}}$, or $\mathcal{D}_F^{\mathbb{C}}$, then we may order $\mathcal{D}_F^{\mathbb{K}}$ as

$$\mathcal{D}^{\mathbb{K}} = \{\mathcal{D}(I_{\alpha_1}), \ldots, \mathcal{D}(I_{\alpha_\rho}); \mathcal{D}(I_{\beta_1}), \ldots, \mathcal{D}(I_{\beta_q}); \ldots; \mathcal{D}(I_{\omega_1}), \ldots, \mathcal{D}(I_{\omega_r})\} \quad (6.66)$$

where $\sigma_{\alpha_1} = \ldots = \sigma_{\alpha_\rho} < \sigma_{\beta_1} = \ldots = \sigma_{\beta_q} < \ldots < \sigma_{\omega_1} = \ldots = \sigma_{\omega_r}$ and $I_{\alpha_1} < \ldots < I_{\alpha_\rho}, I_{\beta_1} < \ldots \ldots < I_{\beta_q}, \ldots, I_{\omega_1} < \ldots < I_{\omega_r}$. Such an ordering for $\mathcal{D}_F^{\mathbb{R}}, \mathcal{D}_F^{\mathbb{C}}$ according to the ordering of the corresponding index sets $\{I_\tau\}, \{I_{\tau'}\}$ will be referred to as natural ordering. The set $\mathcal{D}_F = \{\mathcal{D}_F^{\mathbb{R}}; \mathcal{D}_F^{\mathbb{C}}\}$ for which $\mathcal{D}_F^{\mathbb{R}}$ and $\mathcal{D}_F^{\mathbb{C}}$ have been ordered as above will be called naturally ordered. Note that this ordering is not complete since the elements of the constituent subsets $\mathcal{D}^{\mathbb{R}}(I_\tau), \mathcal{D}^{\mathbb{C}}(I_{\tau'})$ have not been ordered yet. We may now extend Definition (6.1) to the case of Smith-type tuples.

Definition (6.7): Let $F \in S\mathbb{R}_J\{\Theta\}$ and let $\mathcal{D}_F = \{\mathcal{D}_F^{\mathbb{R}}; \mathcal{D}_F^{\mathbb{C}}\}$ be the associated set of e.d. of F over $\mathbb{C}$, which is assumed to be naturally ordered.

(i) For every $\mathcal{D}^{\mathbb{C}}(I_{\tau'}) \in \mathcal{D}_F^{\mathbb{C}}$ we define the set $B'(I_{\tau'})$,

$$B'(I_{\tau'}) \triangleq \{(\gamma_j^{\tau'}, \delta_j^{\tau'}; I_{\tau'}), j \in \underset{\sim}{\nu}_{\tau'}; I_{\tau'} = \{d_1^{\tau'}, \ldots, d_{\sigma_{\tau'}}^{\tau'}\}\} \quad (6.67)$$

as the representation of $\mathcal{D}^{\mathbb{C}}(I_{\tau'})$. Similarly, for every $\mathcal{D}^{\mathbb{R}}(I_\tau) \in \mathcal{D}_F^{\mathbb{R}}$ we define by

$$B(I_\tau) \triangleq \{(\alpha_j^\tau, \beta_j^\tau; I_\tau), j \in \underset{\sim}{\nu}_\tau; I_\tau = \{d_1^\tau, \ldots, d_{\sigma_\tau}^\tau\}\} \quad (6.68)$$

as the representation of $\mathcal{D}^{\mathbb{R}}(I_\tau)$.

(ii) A permutation $\pi\mathcal{B}(\mathcal{I}_\tau),\pi'\mathcal{B}'(\mathcal{I}_{\tau'})$ of the elements of $\mathcal{B}(\mathcal{I}_\tau),\mathcal{B}(\mathcal{I}_{\tau'})$ respectively defines an ordering of the corresponding sets; the set of all possible permutations defined on $\mathcal{B}(\mathcal{I}_\tau),\mathcal{B}'(\mathcal{I}_{\tau'})$ respectively will be denoted by $<\mathcal{B}(\mathcal{I}_\tau)>,<\mathcal{B}'(\mathcal{I}_{\tau'})>$.

(iii) The naturally ordered sets defined from $\mathcal{D}_F^{\mathbb{C}},\mathcal{D}_F^{\mathbb{R}}$ by

$$\mathcal{B}_{\mathbb{C}}(F) \triangleq \{\mathcal{B}'(\mathcal{I}_{\alpha_1'}),\ldots,\mathcal{B}'(\mathcal{I}_{\alpha_{\rho'}'});\ldots;\mathcal{B}'(\mathcal{I}_{\omega_1'}),\ldots,\mathcal{B}'(\mathcal{I}_{\omega_{r'}'})\} \tag{6.69}$$

$$\mathcal{B}_{\mathbb{R}}(F) \triangleq \{\mathcal{B}(\mathcal{I}_{\alpha_1}),\ldots,\mathcal{B}(\mathcal{I}_{\alpha_\rho});\ldots;\mathcal{B}(\mathcal{I}_{\omega_1}),\ldots,\mathcal{B}(\mathcal{I}_{\omega_r})\} \tag{6.70}$$

characterise the sets of real and complex conjugate e.d. of $\mathcal{D}_F$ and shall be called the complex, real unique factorisation sets of F respectively. The sets of integers characterising the possible sets of degrees and the corresponding multiplicities, of e.d. $\mathcal{J}_{\mathbb{C}}(F)$, $\mathcal{J}_{\mathbb{R}}(F)$, where

$$\mathcal{J}_{\mathbb{C}}(F) \triangleq \{(\mathcal{I}_{\alpha_1'},\nu_{\alpha_1'}),\ldots,(\mathcal{I}_{\alpha_{\rho'}'},\nu_{\alpha_{\rho'}'});\ldots;(\mathcal{I}_{\omega_1'},\nu_{\omega_1'}),\ldots,(\mathcal{I}_{\omega_{r'}'},\nu_{\omega_{r'}'})\} \tag{6.71}$$

$$\mathcal{J}_{\mathbb{R}}(F) \triangleq \{(\mathcal{I}_{\alpha_1},\nu_{\alpha_1}),\ldots,(\mathcal{I}_{\alpha_\rho},\nu_{\alpha_\rho});\ldots;(\mathcal{I}_{\omega_1},\nu_{\omega_1}),\ldots,(\mathcal{I}_{\omega_r},\nu_{\omega_r})\} \tag{6.72}$$

are defined as the complex list, real list of F respectively. The sets $\mathcal{B}(F) \triangleq \{\mathcal{B}_{\mathbb{R}}(F);\mathcal{B}_{\mathbb{C}}(F)\}$ and $\mathcal{J}(F) \triangleq \{\mathcal{J}_{\mathbb{R}}(F);\mathcal{J}_{\mathbb{C}}(F)\}$ are defined as the unique factorisation set (UFS) and the list of F correspondingly.

(iv) Every permutation of the elements of $\mathcal{B}(F)$ defined by

$$\bar{\pi}(\mathcal{B}(F)) \triangleq \{\pi(\mathcal{B}(\mathcal{I}_{\alpha_1})),\ldots,\pi(\mathcal{B}(\mathcal{I}_{\omega_r}));\pi'(\mathcal{B}'(\mathcal{I}_{\alpha_1'})),\ldots,\pi'(\mathcal{B}'(\mathcal{I}_{\omega_{r'}'})\} \tag{6.73}$$

where $\pi(\mathcal{B}(\mathcal{I}_\tau))\in<\mathcal{B}(\mathcal{I}_\tau)>,\pi'(\mathcal{B}'(\mathcal{I}_{\tau'}))\in<\mathcal{B}'(\mathcal{I}_{\tau'})>$ defines a complete natural ordering of $\mathcal{B}(F)$; the set of all such permutations will be denoted by $<\mathcal{B}(F)>$

(v) Let $\bar{\pi}(\mathcal{B}(F))=\{\ldots,\pi(\mathcal{B}(\mathcal{I}_\tau)),\ldots;\ldots,\pi'(\mathcal{B}'(\mathcal{I}_{\tau'})),\ldots\}\in<\mathcal{B}(F)>$, where

$$\pi(\mathcal{B}(\mathcal{I}_\tau)) = \{(\alpha_1^\tau,\beta_1^\tau;\mathcal{I}_\tau),\ldots,(\alpha_{\nu_\tau}^\tau,\beta_{\nu_\tau}^\tau;\mathcal{I}_\tau)\}\in<\mathcal{B}(\mathcal{I}_\tau)> \tag{6.74}$$

$$\pi'(\mathcal{B}'(\mathcal{I}_{\tau'})) = \{(\gamma_1^{\tau'},\delta_1^{\tau'};\mathcal{I}_{\tau'}),\ldots,(\gamma_{\nu_{\tau'}}^{\tau'},\delta_{\nu_{\tau'}}^{\tau'};\mathcal{I}_{\tau'})\}\in<\mathcal{B}'(\mathcal{I}_{\tau'})> \tag{6.75}$$

A matrix representation of $\bar{\pi}(\mathcal{B}(F))$ may be defined by

$$T_F^{\pi,\pi'}=\left[\begin{array}{c}\vdots\\ [\mathcal{B}^\pi(\mathcal{I}_\tau)]\\ \vdots\\ \hline \vdots\\ [\mathcal{B}'^{\pi'}(\mathcal{I}_{\tau'})]\\ \vdots\end{array}\right]=\left[\begin{array}{c}[\mathcal{B}_{\mathbb{R}}^\pi(F)]\\ \hline [\mathcal{B}_{\mathbb{C}}^{\pi'}(F)]\end{array}\right],[\mathcal{B}^\pi(\mathcal{I}_\tau)]=\begin{bmatrix}\alpha_1^\tau & \beta_1^\tau\\ \vdots & \vdots\\ \alpha_{\nu_\tau}^\tau & \beta_{\nu_\tau}^\tau\end{bmatrix},[\mathcal{B}'^{\pi'}(\mathcal{I}_{\tau'})]=\begin{bmatrix}\gamma_1^{\tau'} & \delta_1^{\tau'}\\ \vdots & \vdots\\ \gamma_{\nu_{\tau'}}^{\tau'} & \delta_{\nu_{\tau'}}^{\tau'}\end{bmatrix} \tag{6.76}$$

The matrices $T_F^{\pi,\pi'}$, $T_{\mathbb{R},F}^{\pi}\triangleq[\mathcal{B}_{\mathbb{R}}^\pi(F)]$, $T_{\mathbb{C},F}^{\pi'}\triangleq[\mathcal{B}_{\mathbb{C}}^{\pi'}(F)]$ will be referred to as a <u>(π,π')-</u>, <u>$(\mathbb{R},\pi)$-</u>, <u>$(\mathbb{C},\pi')$-basis matrix</u> of F respectively. □

From the above definition it is clear that the results presented for E_p-equivalence of a single polynomial may be naturally extended to the case of E_p-equivalence defined on Smith-type tuples. In the following it will be assumed that F is naturally ordered.

<u>Proposition (6.19)</u>: Let $F_1,F_2\in S\mathbb{R}_J\{\Theta\}$ and let $\mathcal{B}(F_i)=\{\mathcal{B}_{\mathbb{R}}(F_i);\mathcal{B}_{\mathbb{C}}(F_i)\}$, $J(F_i)=\{J_{\mathbb{R}}(F_i);J_{\mathbb{C}}(F_i)\}$, i=1,2, be the associated UFS and list of F_i. Necessary and sufficient condition for $F_1E_pF_2$ are that the following conditions hold true:

(i) $J(F_1)=J(F_2) \iff J_{\mathbb{R}}(F_1)=J_{\mathbb{R}}(F_2)$ and $J_{\mathbb{C}}(F_1)=J_{\mathbb{C}}(F_2)$.

(ii) There exist permutations $\bar{\pi}_i(\mathcal{B}(F_i))=(\pi_i(\mathcal{B}_{\mathbb{R}}(F_i)),\pi_i'(\mathcal{B}_{\mathbb{C}}(F_i)))\in<\mathcal{B}(F_i)>$, i=1,2, such that for the corresponding (π_i,π_i')-basis matrices $T_{F_1}^{\pi_1,\pi_1'}$, $T_{F_2}^{\pi_2,\pi_2'}$ we have $T_{F_1}^{\pi_1,\pi_1'}E_{eh}^{r}T_{F_2}^{\pi_2,\pi_2'}$. □

The proof of this result is similar to that given for E_p-equivalence defined on $\mathbb{R}_d\{\Theta\}$. The importance of Proposition (6.19) is that it demonstrates the fact that all results derived for E_p-equivalence on $\mathbb{R}_d\{\Theta\}$ also carry out to the case of E_p-equivalence defined on the set $S\mathbb{R}_J\{\Theta\}$.

The only difference between the two cases is that the sets $\mathcal{B}(F)=\{\mathcal{B}_{\mathbb{R}}(F);\mathcal{B}_{\mathbb{C}}(F)\}$ and $\mathcal{J}(F)=\{\mathcal{J}_{\mathbb{R}}(F);\mathcal{J}_{\mathbb{C}}(F)\}$ are defined in a more general way by Definition (6.7), rather than Definition (6.1) used for $\mathcal{B}(f)$ and $\mathcal{J}(f)$. The invariants of $E_{\mathcal{P}}(F)$ are the same with those of $E_{\mathcal{P}}(f)$ and thus the same definition and names will be used.

With the study of $E_{\mathcal{P}}$-equivalence on $S\mathbb{R}_{J}\{\Theta\}$ completed, we may now proceed to the solution of the original problem, i.e. the study of complete invariants of the $E_{H-\mathcal{B}}(F,G)$ orbit.

6.7 A complete set of invariants of matrix pencils under bilinear strict equivalence

The starting point in our attempt to characterise $E_{H-\mathcal{B}}$-equivalence of matrix pencils is Theorem (6.1); by this theorem the sets of c.m.i. and r.m.i., $\mathcal{I}_c(F,G),\mathcal{I}_r(F,G)$ of $sF-\hat{s}G$ are invariant; and the extra invariants needed to form a complete set are provided by the invariants of the set $\mathcal{F}(F,G)$, of homogeneous invariant polynomials of $sF-\hat{s}G$, under $E_{\mathcal{P}}$-equivalence. Note that $\mathcal{F}(F,G)\in S\mathbb{R}_{J}\{\Theta\}$, and thus the results of the previous section carry over naturally to the case of $\mathcal{F}(F,G)$. Before we state the main result of this section we introduce some notation.

For the set of homogeneous invariant polynomials $\mathcal{F}(F,G)$ of $L(s,\hat{s})=sF-\hat{s}G$ we shall denote by $\mathcal{D}(F,G)$ the symmetric set of e.d. over $\mathbb{C}$, and by $\mathcal{B}(F,G)=\{\mathcal{B}_{\mathbb{R}}(F,G);\mathcal{B}_{\mathbb{C}}(F,G)\}$, $\mathcal{J}(F,G)=\{\mathcal{J}_{\mathbb{R}}(F,G);\mathcal{J}_{\mathbb{C}}(F,G)\}$ the UFS and the list of $\mathcal{F}(F,G)$ and thus of $L(s,\hat{s})$. The set $\mathcal{B}(F,G)$ will be assumed to be naturally ordered and if $\bar{\pi}\mathcal{B}(F,G)=(\pi\mathcal{B}_{\mathbb{R}}(F,G),\pi'\mathcal{B}_{\mathbb{C}}(F,G))\in<\mathcal{B}(F,G)>$, then the (π,π')-basis matrix will be designated by $T_{\pi,\pi'}$, and $\mathcal{T}(F,G)$ will denote the family of matrix representations of $\mathcal{B}(F,G)$ and thus of $L(s,\hat{s})$. If $\mu=\#\mathcal{B}_{\mathbb{R}}(F,G)$, $\nu=\#\mathcal{B}_{\mathbb{C}}(F,G)$, then (μ,ν) will be called the order of $L(s,\hat{s})$. Following Definition (6.6), we have that if $\mu+\nu\geq 3$, then for $\forall\ T_{\pi,\pi'}\in\mathcal{T}(F,G), r\in<\mu+\nu>$, $(i_1,i_2)\in Q^{r)}_{2,\mu+\nu-1}$, $\tilde{\underline{g}}^{\pi,\pi'}_{r,i_1,i_2}$ will denote the (r,i_1,i_2)-$\mathbb{R}$-CGV of $T_{\pi,\pi'}$, and shall be called the (π,π')-(r,i_1,i_2)-$\mathbb{R}$-canonical Grassmann vector of $L(s,\hat{s})$;

the set of all such vectors (eqn(6.58)), is designated by $\mathcal{G}(F,G)$ and we shall call it the $\mathbb{R}$-canonical Grassmann vector set of $L(s,\hat{s})$. For every $T_{\pi,\pi'}\in\mathcal{T}(F,G)$ we shall denote by $\mathcal{P}_r^{\pi,\pi'}(F,G)$, $\mathcal{P}^{\pi,\pi'}(F,G)$ and by $\mathcal{P}^*(F,G)$ the set of (π,π')-r-prime Plücker vectors ($r\in<\mu+\nu>$) of $T_{\pi,\pi'}$, the set of all r-prime Plücker vectors of $T_{\pi,\pi'}$, and the Plücker vector set of $\mathcal{T}(F,G)$ and thus of $L(s,\hat{s})$ (see eqn(6.62)). If $(\mu,\nu)=(0,2)$, the α-, β-, $\bar{\alpha}$-, $\bar{\beta}$-characteristic matrices of $L(s,\hat{s})$ are defined as for the case of polynomials (eqn(6.59),(6.60)). Finally, $L(s,\hat{s})\in\mathcal{L}(\Theta)$ will be called a general order pencil if $(\mu,\nu)\neq\{(1,0),(2,0),(3,0),(0,1),(1,1),(0,2)\}$. We may now state the main result of this chapter.

Theorem (6.7): Let $L(s,\hat{s})=sF-\hat{s}G\in\mathcal{L}(\Theta)$ be a general order matrix pencil and let $\mathcal{B}(F,G)$, $\mathcal{J}(F,G)=\{\mathcal{J}_{\mathbb{R}}(F,G);\mathcal{J}_{\mathbb{C}}(F,G)\}$, $\mathcal{G}(F,G)$, $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$ be the sets associated with $L(s,\hat{s})$. A complete set of invariants for the $\mathcal{E}_{\mathcal{H}\text{-}\mathcal{B}}(F,G)$ equivalence class is defined by:

(i) $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$.

(ii) $\mathcal{J}(F,G)=\{\mathcal{J}_{\mathbb{R}}(F,G);\mathcal{J}_{\mathbb{C}}(F,G)\}$.

(iii) $\mathcal{G}(F,G)$, or $\mathcal{G}(F,G)$ modulo $-\mathcal{C}_2(\bar{E})$, where $\bar{E}=\text{diag}\{E,\dots,E\}$, $E=[j]_{\mathbb{R}}$. □

The proof of this result readily follows from Theorems (6.1),(6.6) and Proposition (6.19). $\mathcal{E}_{\mathcal{H}\text{-}\mathcal{B}}$-equivalence for a number of special type matrix pencils is treated by the following corollaries.

Corollary (6.11): Let $L(s,\hat{s})=sF-\hat{s}G\in\mathcal{L}(\Theta)$ and let the order (μ,ν) of $L(s,\hat{s})$ take values from the set $\{(1,0),(2,0),(3,0),(0,1),(1,1)\}$. A complete set of invariants for the $\mathcal{E}_{\mathcal{H}\text{-}\mathcal{B}}(F,G)$ orbit is defined by the sets: $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$, and $\mathcal{J}(F,G)=\{\mathcal{J}_{\mathbb{R}}(F,G),\mathcal{J}_{\mathbb{C}}(F,G)\}$. □

Corollary (6.12): Let $L_1(s,\hat{s})=sF_1-\hat{s}G_1$, $L_2(\lambda,\hat{\lambda})\in\mathcal{L}(\Theta)$ be two pencils of order $(\mu,\nu)=(0,2)$. Let $\mathcal{J}_{\mathbb{C}}(F_i,G_i)=\{(\mathcal{I}_{\tau_i},1),(\mathcal{I}_{\tau_i'},1)\}$, $\{P_i^{\alpha},P_i^{\beta},P_i^{\bar{\alpha}},P_i^{\bar{\beta}}\}$, $\mathcal{I}_c(F_i,G_i)$ and $\mathcal{I}_r(F_i,G_i)$, i=1,2, be the corresponding lists, characteristic matrices and minimal indices sets respectively of the two pencils. $L_1(s,\hat{s})\mathcal{E}_{H\text{-}B}L_2(\lambda,\hat{\lambda})$, if and only if the following conditions hold true:

(i) $\mathcal{I}_c(F_1,G_1)=\mathcal{I}_c(F_2,G_2),\mathcal{I}_r(F_1,G_1)=\mathcal{I}_r(F_2,G_2),\mathcal{J}_{\mathbb{C}}(F_1,G_1)=\mathcal{J}_{\mathbb{C}}(F_2,G_2)$.

(ii) (α) If $\mathcal{I}_{\tau_1}<$, or $\dot{<}\mathcal{I}_{\tau'_1}$, then $P_1^{\beta}\square P_2^{\alpha}=0$.

(β) If $\mathcal{I}_{\tau_1}=\mathcal{I}_{\tau_1}$, then either $P_1^{\beta}\square P_2^{\alpha}=0$, or $P_1^{\bar{\beta}}\square P_2^{\alpha}=0$. □

<u>Corollary (6.13)</u>: Let $L(s,\hat{s})=sF-\hat{s}G\in\mathcal{L}(\Theta)$ and let $(\mu,\nu)=(k,0),k\geq4$. Let $\mathcal{I}_c(F,G),\mathcal{I}_r(F,G),\mathcal{J}(F,G)=\mathcal{J}_{\mathbb{R}}(F,G),\mathcal{P}^*(F,G),\mathcal{G}(F,G)$ be the sets associated with $L(s,\hat{s})$. A complete set of invariants for the $\mathcal{E}_{H\text{-}B}(F,G)$-equivalence class is defined by:

(i) $\mathcal{I}_c(F,G),\mathcal{I}_r(F,G),\mathcal{J}_{\mathbb{R}}(F,G)$.

(ii) $\mathcal{G}(F,G)$ modulo ±1, or equivalently $\mathcal{P}^*(F,G)$ modulo collinearity of the corresponding vectors. □

These corollaries follow immediately by Theorem (6.1), Proposition (6.19) and the corresponding results for binary polynomials stated in the previous section. The properties of the invariant sets $\mathcal{G}(F,G)$, or $\mathcal{P}^*(F,G)$, have been studied in detail in Section (6.5), where $\mathcal{E}^r_{eh}$-equivalence was discussed. The set $\mathcal{G}(F,G)$ is generated by a single element (in the sense discussed before) and thus the elements in $\mathcal{G}(F,G)$ are not independent invariants but dependent; however, the whole set $\mathcal{G}(F,G)$ (or $\mathcal{P}^*(F,G)$) is used in the above results, to avoid the use of permutations. Finally, we should note that two sets $\mathcal{G}(F_1,G_1),\mathcal{G}(F_2,G_2)$ are equal if and only if there exist $\tilde{\underline{g}}(1)^{\pi_1,\pi'_1}_{r,i_1,i_2}\in\mathcal{G}(F_1,G_1)$ and $\tilde{\underline{g}}(2)^{\pi_2,\pi'_2}_{q,j_1,j_2}\in\mathcal{G}(F_2,G_2)$ such that

$$\tilde{\underline{g}}(2)^{\pi_2,\pi'_2}_{q,j_1,j_2}=\tilde{\underline{g}}(1)^{\pi_1,\pi'_1}_{r,i_1,i_2},\text{ or }=-\mathcal{C}_2(E)\tilde{\underline{g}}(1)^{\pi_1,\pi'_1}_{r,i_1,i_2}\tag{6.77}$$

Similarly, two sets $\mathcal{P}^*(F_1,G_1),\mathcal{P}^*(F_2,G_2)$ are equal if and only if there exist $\mathcal{P}_r^{\pi_1,\pi'_1}(F_1,G_1)\in\mathcal{P}^*(F_1,G_1)$ and $\mathcal{P}_q^{\pi_2,\pi'_2}(F_2,G_2)\in\mathcal{P}^*(F_2,G_2)$ such that

$$\mathcal{P}_r^{\pi_1,\pi'_1}(F_1,G_1)\nearrow\!\!\nearrow\mathcal{P}_q^{\pi_2,\pi'_2}(F_2,G_2)\tag{6.78}$$

where "$\nearrow\!\nearrow$" stands for collinearity of the corresponding vector sets. The last two conditions express the property described by Remarks (6.16),(6.17) respectively as "equality of a pair of representative points".

Remark (6.19): If $L(s,\hat{s})=sF-\hat{s}G\in\mathbb{R}^{m\times n}[s,\hat{s}],m\neq n$, is a generic pencil ($L(s,\hat{s})$ full rank and $\mathcal{S}^*(s,\hat{s})$ has no nontrivial elements, i.e. $\mathcal{F}(F,G)=\emptyset$), then a complete set of invariants under $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence is defined by:

(i) $\mathcal{I}_c(F,G)$, if $m<n$.

(ii) $\mathcal{I}_r(F,G)$, if $m>n$.

For such pencils, the notions of $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ and $\mathcal{E}_{\mathcal{H}}$ equivalence coincide; since for both types of equivalence the same set is a complete invariant. Thus, if we keep the same pair of indeterminates $(s,\hat{s})$, then $\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)=\mathcal{E}_{\mathcal{H}}(F,G)$. □

For regular pencils the notions of $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ and $\mathcal{E}_{\mathcal{H}}$ equivalence are quite distinct, since there is no special case where the two notions may coincide.

Testing for $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ equivalence of two pencils $L_1(s,\hat{s})=sF_1-\hat{s}G_1$, $L_2(\lambda,\hat{\lambda})=$ $=\lambda F_2-\hat{\lambda}G_2\in\mathcal{L}$ is an elaborate procedure, since finally the sets $\mathcal{G}(F_i,G_i)$, or $\mathcal{P}^*(F_i,G_i)$ have to be computed. Note that the rank ρ of a pencil and the set of indices $\mathcal{J}=\{d_i:i\in\underset{\sim}{g},d_i=\deg f_i(s,\hat{s}),f_i(s,\hat{s})\in\mathcal{F}(F,G)\}$ are invariants under $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence, but not independent from the sets defined before. A systematic procedure for testing $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence implies the following steps:

(i) Find the ranks ρ_1,ρ_2 of $L_1(s,\hat{s}),L_2(\lambda,\hat{\lambda})$. If $\rho_1\neq\rho_2$ we stop, since $L_1(s,\hat{s})\not{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$.

(ii) If $\rho_1=\rho_2$, find the sets of degrees $\mathcal{J}_1,\mathcal{J}_2$ of $\mathcal{F}(F_i,G_i)$, i=1,2. If $\mathcal{J}_1\neq\mathcal{J}_2$ stop, since then $L_1(s,\hat{s})\not{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$.

(iii) If $\mathcal{J}_1=\mathcal{J}_2$, compute $\mathcal{J}(F_i,G_i)$, i=1,2. If $\mathcal{J}(F_1,G_1)\neq\mathcal{J}(F_2,G_2)$ stop, since then $L_1(s,\hat{s})\not{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$.

(iv) If $\mathcal{J}(F_1,G_1)=\mathcal{J}(F_2,G_2)$, compute $\mathcal{I}_c(F_i,G_i),\mathcal{I}_r(F_i,G_i)$, i=1,2. If $\mathcal{I}_c(F_1,G_1)\neq\mathcal{I}_c(F_2,G_2)$, or $\mathcal{I}_r(F_1,G_1)\neq\mathcal{I}_r(F_2,G_2)$ stop, since then

$L_1(s,\hat{s})\not{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$.

(v) If $\mathcal{I}_c(F_1,G_1)=\mathcal{I}_c(F_2,G_2)$ and $\mathcal{I}_r(F_1,G_1)=\mathcal{I}_r(F_2,G_2)$ then compute the UFS $\mathcal{B}(F_i,G_i)=\{\mathcal{B}_{\mathbb{R}}(F_i,G_i);\mathcal{B}_{\mathbb{C}}(F_i,G_i)\}$, i=1,2. The sets $\mathcal{B}_{\mathbb{R}}(F_i,G_i),\mathcal{B}_{\mathbb{C}}(F_i,G_i)$ characterise the sets or real and complex e.d. $\mathcal{D}_{\mathbb{R}}(F_i,G_i),\mathcal{D}_{\mathbb{C}}(F_i,G_i)$ respectively and thus define sets of homogeneous polynomials $\mathcal{F}_{\mathbb{R}}(F_i,G_i)\in S\mathbb{R}_{J_\alpha}\{\Theta\}$, $\mathcal{F}_{\mathbb{C}}(F_i,G_i)\in S\mathbb{R}_{J_\beta}\{\Theta\}$. If $\mathcal{F}_{\mathbb{R}}(F_1,G_1)\not{E}_{\mathcal{P}}\mathcal{F}_{\mathbb{R}}(F_2,G_2)$, or $\mathcal{F}_{\mathbb{C}}(F_1,G_1)\not{E}_{\mathcal{P}}\mathcal{F}_{\mathbb{C}}(F_2,G_2)$, then stop, since $\mathcal{F}(F_1,G_2)\not{E}_{\mathcal{P}}\ (F_2,G_2)$ and thus $L_1(s,\hat{s})\not{E}_{\mathcal{H}-\mathcal{B}}L_2(\lambda,\hat{\lambda})$.

(vi) If $\mathcal{F}_{\mathbb{R}}(F_1,G_1)E_{\mathcal{P}}\mathcal{F}_{\mathbb{R}}(F_2,G_2)$ and $\mathcal{F}_{\mathbb{C}}(F_2,G_2)E_{\mathcal{P}}(F_2,G_2)$, then proceed to the computation of $\mathcal{G}(F_i,G_i)$, i=1,2, or the special cases tests for checking the $E_{\mathcal{P}}$-equivalence of $\mathcal{F}(F_i,G_i)$, i=1,2.

The $E_{\mathcal{H}-\mathcal{B}}(F,G)$ equivalence class has been characterised by a complete set of invariants; this set of invariants is common to all pencils $L_\theta,\tilde{L}_{\tilde{\theta}}\in$ $\in E_{\mathcal{H}-\mathcal{B}}(F,G)$, but there exist a number of other functions, defined on a pencil, which are not $E_{\mathcal{H}-\mathcal{B}}$ invariant and thus generally they take different values on the elements $L_\theta,\tilde{L}_{\tilde{\theta}}$ of the class. In a number of cases, it is desirable to find a matrix pencil $L_\theta\in E_{\mathcal{H}-\mathcal{B}}$ with a prescribed set of characteristics; such a problem may be of interest in the study of numerical analysis aspects, or system theoretic applications of the matrix pencil theory. The following problem then arises:

Problem: Let $L_\theta=sF-\hat{s}G\in\mathcal{L}(\Theta)$, $E_{\mathcal{H}-\mathcal{B}}(F,G)$ the corresponding equivalence class, $\mathcal{R}$ the set of complete invariants of $E_{\mathcal{H}-\mathcal{B}}(F,G)$ (invariant functions defined on $\forall\ L'_{\theta'}\in E_{\mathcal{H}-\mathcal{B}}(F,G)$) and let $\mathcal{K}$ be a set of functions defined on $\forall\ L'_{\theta'}\in E_{\mathcal{H}-\mathcal{B}}(F,G)$ for which $\mathcal{K}\cap\mathcal{R}=\emptyset$. Let us further denote by $\mathcal{K}(F',G')$ the set of values of $\mathcal{K}$ on the pencil $L_{\theta'}=\lambda F'-\hat{\lambda}G'$. Find whether there exist $b\in\mathcal{B}$ and $h\in\mathcal{H}$, such that $\tilde{L}_{\tilde{\theta}}=(h*b)\circ L_\theta=\lambda\tilde{F}-\hat{\lambda}\tilde{G}$ has a prescribed set of values $\mathcal{K}(\tilde{F},\tilde{G})=A_o$. This problem will be referred to as a $\mathcal{K}$-characteristics $E_{\mathcal{H}-\mathcal{B}}$ pencil assignment problem, and shall be denoted in short by $\mathcal{K}$-$E_{\mathcal{H}-\mathcal{B}}$-PA.

The nature of the particular $\mathcal{K}$-$E_{\mathcal{H}-\mathcal{B}}$-PA problem depends on the type of characteristic functions K we specify. We close this chapter by discussing a special type of K-E_{H-B}-PA problem related to the stability of the e.d. of the pencil.

Definition (6.8): Let $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$. $f(s,\hat{s})$ will be said to be stable if the polynomials $f(s,1)$ and $f(1,\hat{s})$ have no roots in the closed right-half $\mathbb{C}$-plane. A pencil $L_\theta=sF-\hat{s}G$ will be called stable if the set $F(F,G)$ is stable.

Remark (6.20): Let $f(s,\hat{s})\in\mathbb{R}_d\{\Theta\}$ and let $M(f)$ be the set of e.d. of $f(s,\hat{s})$ over $\mathbb{R}$, that is $M(f)=\{(\alpha_i s-\beta_i\hat{s})^{k_i},(r_j s^2+p_j s\hat{s}+q_j\hat{s}^2)^{\tau_j},i\in\underset{\sim}{\mu},j\in\underset{\sim}{\nu}\}$. $f(s,\hat{s})$ will be stable if and only if $\alpha_i\beta_i<0$ for $\forall$ $i\in\underset{\sim}{\mu}$ and $r_j,p_j,q_j>0$ (or <0) for $\forall$ $j\in\underset{\sim}{\nu}$.

A stable polynomial $f(s,\hat{s})$ has no e.d. of the type $s^p,\hat{s}^q$ and all of the finite nonzero roots are in the open right half $\mathbb{C}$-plane. The problem considered next may be stated as follows:

Stabilizability of e.d. problem (SEDP): Determine whether there exists $\tilde{L}_{\tilde{\theta}}=(\lambda\tilde{F}-\hat{\lambda}\tilde{G})\in E_{H-B}(F,G)$, such that $\tilde{L}_{\tilde{\theta}}$ is stable. For this problem the set K is the $F(\tilde{F},\tilde{G})$ and the prescribed property is the stability of the elements of $F(\tilde{F},\tilde{G})$.

Before we proceed with the study of solvability of SEDP we define the following subset of $PGL(1,\mathbb{C}/\mathbb{R})$:

$$PGL_+(1,\mathbb{C}/\mathbb{R})\triangleq\{\beta_+:\begin{bmatrix}s\\\hat{s}\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}\lambda\\\hat{\lambda}\end{bmatrix},a,b,c,d>0,\ ad-cb\neq0\ \}\qquad(6.79)$$

Projective transformations of the type β_+, defined as above, will be called positive-real projective transformations and the transformations $b_+\in B$ induced by a $\beta\in PGL_+(1,\mathbb{C}/\mathbb{R})$ will be referred with the same name. The subset of B containing all positive real transformations will be denoted by B_+.

<u>Proposition (6.20)</u>: Let $f(s,\hat{s})=(\alpha s-\beta\hat{s})^k$, $g(s,\hat{s})=(rs^2+ps\hat{s}+q\hat{s}^2)^\tau$ be two irreducible over $\mathbb{R}$ binary polynomials.

(i) If $f(s,\hat{s}),g(s,\hat{s})$ are stable, then for $\forall\ \beta\epsilon PGL_+(1,\mathbb{C}/\mathbb{R})$ the polynomials $\beta\circ f(s,\hat{s})=\tilde{f}(\lambda,\hat{\lambda}),\beta\circ g(s,\hat{s})=\tilde{g}(\lambda,\hat{\lambda})$ are stable.

(ii) (a) If $f(s,\hat{s})$ is unstable, then necessary and sufficient condition for $\beta\circ f(s,\hat{s})=\tilde{f}(\lambda,\hat{\lambda}),\beta\epsilon PGL_+(1,\mathbb{C}/\mathbb{R})$ to be stable is that

$$\alpha/\beta > \max\{c/a,d/b\} \tag{6.80}$$

(b) If $g(s,\hat{s})$ is unstable, then necessary and sufficient condition for $\beta\circ g(s,\hat{s})=\tilde{g}(\lambda,\hat{\lambda}),\beta\epsilon PGL_+(1,\mathbb{C}/\mathbb{R})$ to be stable is that

$$-p/r < 2ab/(ad+cb) + q/r.2cd/(ad+cb) \tag{6.81}$$

<u>Proof</u>

(i) The coefficients of $\tilde{f}(\lambda,\hat{\lambda}),\tilde{g}(\lambda,\hat{\lambda})$ are defined by

$$[\tilde{\alpha},-\tilde{\beta}]=[\alpha,-\beta]\begin{bmatrix} a & b \\ c & d \end{bmatrix} = [\alpha a-\beta c,\alpha b-\beta d] \tag{6.82a}$$

$$[\tilde{r},\tilde{p},\tilde{q}] = [r,p,q]\begin{bmatrix} a^2 & 2ab & b^2 \\ ac & ad+cb & bd \\ c^2 & 2cd & d^2 \end{bmatrix} \tag{6.82b}$$

Clearly, if $\alpha,-\beta,r,p,q>0$, then the resulting coefficients $\tilde{\alpha},-\tilde{\beta},\tilde{r},\tilde{p},\tilde{q}$ of $\tilde{f}(\lambda,\hat{\lambda}),\tilde{g}(\lambda,\hat{\lambda})$ are positive, if $a,b,c,d>0$, i.e. $\tilde{f}(\lambda,\hat{\lambda}),\tilde{g}(\lambda,\hat{\lambda})$ are stable.

(ii) (a) For $\tilde{f}(\lambda,\hat{\lambda})$ to be stable $\alpha a-\beta c>0$ and $\alpha b-\beta d>0$. Since $\beta>0$, and $a,b>0$, then, $\alpha/\beta>c/a$ and $\alpha/\beta>b/d$ and (6.80) follows. The sufficiency is established by a mere reversion of the arguments.

(b) Since $g(s,\hat{s})$ is irreducible, $\tilde{\Delta}=\tilde{p}^2-4\tilde{r}\tilde{q}<0$ for all $\beta\epsilon PGL(1,\mathbb{C}/\mathbb{R})$; thus, if $\tilde{r}>0$, automatically $\tilde{q}$ is also positive. For stabilization, we should have

$$\tilde{p} = 2rab+p(ad+bc)+2qcd>0$$

Since r,a,b,c,d>0, then the above condition yields (6.81). The sufficiency follows by reversing the steps. □

<u>Remark (6.21)</u>: Every positive real transformation β stabilizes e.d. of the type s^p and $\hat{s}^q$.

The above result shows that positive real projective transformations preserve stability of stable e.d. and under the conditions (6.80),(6.81) may be used to stabilize unstable e.d. Condition (6.81) is rather difficult to handle; a simpler condition is established by the following result.

<u>Corollary (6.14)</u>: Let $g(s,\hat{s})=(rs^2+ps\hat{s}+q\hat{s}^2)^\tau$ be an unstable binary quadratic($r,p>0,p\leq 0$). A transformation $\beta\epsilon PGL_+(1,\mathbb{C}/\mathbb{R})$ exists such that $\beta\circ g(s,\hat{s})=\tilde{g}(\lambda,\hat{\lambda})$ is stable, if either of the following conditions hold true:

$$-p/r < 2ab/(ad+bc) \tag{6.83a}$$

or

$$-p/q < 2cd/(ad+bc) \tag{6.83b}$$

□

The sufficiency conditions presented by (6.83a), or (6.83b) readily follow from condition (6.81). Proposition (6.20) and its corollary may be used to establish the conditions for stabilizability of $F(F,G)$ under a positive real projective transformation. For the set $F(F,G)$ of homogeneous invariant polynomials of $L_\theta=sF-\hat{s}G$, we shall denote by M_ℓ,M_q the sets unstable linear, quadratic e.d. over $\mathbb{R}$ respectively, which exclude e.d. of the type $s^p,\hat{s}^q$. If

$$M_\ell = \{(\alpha_i s-\beta_i\hat{s})^{k_i},\alpha_i\beta_i>0,i\epsilon\underset{\sim}{\mu},k_i\in\mathbb{Z}\}$$

$$M_q = \{(r_j s^2+p_j s\hat{s}+q_i\hat{s}^2)^{\tau_j},r_j p_j<0,j\epsilon\underset{\sim}{\nu},\tau_j\in\mathbb{Z}\}$$

then the numbers defined from M_ℓ,M_q by

$$\lambda_{\min} \stackrel{\Delta}{=} \min\{\alpha_i/\beta_i : i\in\underset{\sim}{\mu}\} \tag{6.84a}$$

$$\rho^1_{\max} \stackrel{\Delta}{=} \max\{-p_j/r_j : j\in\underset{\sim}{\nu}\},\rho^2_{\max} \stackrel{\Delta}{=} \max\{-p_j/q_j, j\in\underset{\sim}{\nu}\} \tag{6.84b}$$

will be referred to as the instability indices of $\mathcal{F}(F,G)$. Using the instability indices $\lambda_{\min},\rho^1_{\max},\rho^2_{\max}$ we may state the following result for stabilizability of L_θ under positive real projective transformations.

Theorem (6.8): Let $(\lambda_{\min},\rho^1_{\max},\rho^2_{\max})$ be the instability indices of $L_\theta=sF-\hat{s}G$. There exists a positive real transformation $b\in\mathcal{B}_+$ such that $b\circ L_\theta=\lambda\tilde{F}-\hat{\lambda}\tilde{G}$ is stable, if there exist a,b,c,d>0 such that either of the following two conditions are satisfied:

(i) $\lambda_{\min}>\max\{c/a,d/b\}$ and $\rho^1_{\max}<2ab/(ad+cb)$,

or

(ii) $\lambda_{\min}>\max\{c/a,d/b\}$ and $\rho^2_{\max}<2ab/(ad+cb)$

Proof

By Proposition (6.20) and Remark (6.21) a positive real projective transformation preserves stability of stable e.d., stabilizes e.d. of the type $s^p,\hat{s}^q$ and stabilizes unstable e.d. of the type $(\alpha s-\beta\hat{s})^k,(rs^2+ps\hat{s}+q\hat{s})^\tau$, if conditions (6.80), (6.83a) or (6.83b) hold true. If $\lambda_{\min}>\max\{c/a,d/b\}$, all e.d. of the type $(\alpha s-\beta\hat{s})^k$ are stabilized. Similarly, by Corollary (6.14), if $\rho^1_{\max}$, or $\rho^2_{\max}<2ab/(ad+cb)$ then all e.d. $(rs^2+ps\hat{s}+q\hat{s}^2)^\tau$ are stabilized. □

Corollary (6.15): There always exist a $b\in\mathcal{B}_+$ such that the pencil $b\circ L_\theta=\lambda\tilde{F}-\hat{\lambda}\tilde{G}\in\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)$ is stable. The parameters (a,b,c,d) of a positive real projective transformation β that induces b are determined by

$$c/a+d/b < \min\{\lambda_{\min},2/\rho^1_{\max}\} \tag{6.85}$$

Proof

By condition (6.85) we have that

$$c/a+d/b < \lambda_{min} \text{ and } c/a+d/b < 2/\rho^{1}_{max} \qquad (6.86)$$

However, $\max\{c/a,d/b\}<c/a+d/b<\lambda_{min}$ and thus the first of conditions of Theorem (6.8) is satisfied. The second of conditions (6.86) is equivalent to the second of the conditions of Theorem (6.8) and this completes the proof.

□

Corollary (6.15) shows that SEDP has always a solution and it also indicates how the parameters of a positive real projective transformation may be chosen. The family of $PGL_{+}(1,\mathbb{C}/\mathbb{R})$ has been chosen in this investigation because of its property to preserve stability of stable e.d. The importance of SEDP for linear systems is a topic discussed in a following section.

6.8 Conclusions

The notion of bilinear-strict equivalence of matrix pencils has been introduced and a complete set of invariants for the $\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)$ orbit has been defined. This work extends the classical results of the Weierstrass-Kronecker theory of strict equivalence and poses a number of new questions for matrix pencil theory such as the study of different types of $\mathcal{K}$-$\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-PA problems and the search for a canonical form under $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence. Of particular importance, from the numerical viewpoint, is the search for $\tilde{L}_{\tilde{\theta}}\in\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)$ which are "well conditioned" in some sense for computations. The study of SEDP provides a partial answer to this problem, since for the case of regular pencils $\tilde{L}_{\tilde{\theta}}=\lambda\tilde{F}-\hat{\lambda}\tilde{G}\in\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)$ may be found with $|\tilde{F}|,|\tilde{G}|\neq 0$ (no zero and infinite e.d.). The most important, from the numerical point of view, problem is of course the assignment of the condition number. This problem is one of the topics left for future research.

The $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence of matrix pencils provides the means for the classification of the system theoretic properties into two classes: those which are frequency-space transformation invariant, and those which are frequency transformation dependent. It is clear from the results of this section, that system properties associated with the sets $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$, $\mathcal{J}(F,G)$ are frequency-space independent, whereas those depending on the nature of the roots of the e.d. are frequency dependent. The complete study of this "relativistic" classification of system properties will be presented in the following section. $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-equivalence also provides us with the tools for defining suitable system duals, which may be used for the investigation of various system theoretic properties.

The definition of a canonical form for the equivalence class $\mathcal{E}_{\mathcal{H}-\mathcal{B}}(F,G)$ and the search for a complete set of invariants under $\mathcal{E}_{\mathcal{B}}$-equivalence, are problems still open for future research. An important by-product of the work in this chapter is the solution of the general linear mapping problem and the study of the $\mathcal{E}^{c}_{eh}$-, $\mathcal{E}^{r}_{eh}$-equivalence. It seems that $\mathcal{E}^{c}_{eh}$-equivalence, appropriately generalized over rings, could be important in the study of problems such as the simultaneous stabilization [Sae & Mur - 1].

CHAPTER 7:

Geometric and Dynamic aspects of the Linear Generalised Autonomous Differential Systems

CHAPTER 7: GEOMETRIC AND DYNAMIC ASPECTS OF THE LINEAR GENERALISED AUTONOMOUS DIFFERENTIAL SYSTEMS

7.1 Introduction

The generalised autonomous differential system $\mathcal{S}(F,G)$: $F\dot{\underline{x}}=G\underline{x}$, $F,G\in\mathbb{R}^{m\times n}$ has emerged as the unifying description to which problems of the regular, extended state space theory of linear systems may be reduced [Kar. & Hay. -1]. The central problem of linear geometric theory [Won. -1], [Will. -1] of regular state space geometric theory is the dynamic and geometric characterisation of the subspaces of the state space; the algebrization of the fundamental tools of regular state space geometric theory in terms of matrix pencils [Kar. -1], [Jaf. & Kar. -1] has demonstrated that the algebraic structure of the restriction pencil (expressed in terms of the strict equivalence invariants) is the key tool from which the geometric and dynamic aspects of regular state space theory may be deduced. The aim of this chapter is to extend the treatment given in [Kar. -1], [Jaf. & Kar. -1] for regular state space systems to the case of $\mathcal{S}(F,G)$; such a study provides the means for a unifying treatment of the geometric and dynamic properties of regular and extended state space theory. The algebraic and number theoretic properties of matrix pencils developed in Chapters (4) and (5) provide the basis for the study of the geometry of the subspaces of the domain of (F,G) as well as the foundations of an algorithm, based on the properties of PAPS, for the computation of the Kronecker canonical form. The notion of the invariant forced realization [Kar. & Hay. -1,2] is further developed and it is shown to be of crucial importance for the dynamic characterisation of the subspaces of the domain of (F,G). In fact, it is shown that problems defined on $\mathcal{S}(F,G)$ may be reduced to equivalent problems of the regular state space theory.

The chapter is structured as follows: In section (7.2) we introduce some notation and we give some results extending the theory of PAPS developed for regular pencils to the general case of singular pencils. These results yield a procedure for the computation of the Kronecker canonical form which is based on two types of PAPS. Section (7.3) deals with the geometry of the subspaces of the pair (F,G). The notion of the restriction pencil $(F,G)/\mathcal{V}$, of a given subspace $\mathcal{V}$ of the domain of (F,G) is introduced and the various subspaces $\mathcal{V}$ are classified in terms of the invariants of the pencil $(F,G)/\mathcal{V}$. Subsequently, the different types of subspaces are characterised in terms of geometric and number theoretic conditions and the notions of (F,G)-invariant, (G,F)-invariant, and (F,G)-completely invariant subspaces are introduced. The new notions introduced to this section are natural extensions of the standard geometric notions, of (A,B)-invariance, almost (A,B)-invariance, of controllability and almost controllability subspaces. In Section (7.4) the notion of invariant forced realizations of $\mathcal{S}(F,G)$ [Kar. & Hay. -2] is discussed and some further results are derived; these results demonstrate that problems of characterisation of subspaces of extended state space systems may always be reduced to equivalent problems of regular state space theory. In Section (7.4) the family of solutions of $\mathcal{S}(F,G)$ for a given initial condition is derived and its properties are discussed. These results demonstrate that a general (non-square) differential system, although it does not define a dynamical system, is closely associated with a dynamical system. In the case where there is no uniqueness of solution, for a given initial condition, the family of solutions is parametrized in terms of external functions of an invariant forced realization. Finally, the various types of subspaces introduced in Section (7.3) are characterised dynamically and the notions of reachability, and system description redundancy for $\mathcal{S}(F,G)$ are defined.

7.2 The Piecewise Arithmetic Progression Sequences (PAPS) of a general pencil and the Kronecker canonical form

Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$, $\mathrm{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}=\rho\le\min(m,n)$ be a general pencil (singular, or regular) and let us denote by $\mathcal{I}_c(F,G), \mathcal{I}_r(F,G), \mathcal{D}_o(F,G), \mathcal{D}_\infty(F,G)$, $\mathcal{D}_\alpha(F,G)$ the sets of c.m.i., r.m.i., zero-e.d., infinite-e.d. and α-e.d., $\alpha\in\mathbb{C}-\{0\}$, of the pencil $sF-\hat{s}G$ respectively. Throughout this chapter we shall adopt the following two equivalent descriptions for the above sets:

$$\mathcal{I}_c(F,G) \triangleq \{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\le\ldots\le\varepsilon_p\}$$

$$\text{OR} \quad \triangleq \{(\varepsilon_i,\rho_i), i\in\underset{\sim}{\mu}, 0\le\varepsilon_1<\ldots<\varepsilon_\mu, p=\sum_{i=1}^{\mu}\rho_i\} \tag{7.1a}$$

$$\mathcal{I}_r(F,G) \triangleq \{\zeta_1=\ldots=\zeta_{g'}=0<\zeta_{g'+1}\le\ldots\le\zeta_t\}$$

$$\text{OR} \quad \triangleq \{(\zeta_i,\pi_i), i\in\underset{\sim}{\mu}', 0\le\zeta_1<\ldots<\zeta_{\mu'}, t=\sum_{i=1}^{\mu'}\pi_i\} \tag{7.1b}$$

$$\mathcal{D}_o(F,G) \triangleq \{s^{p_i}, 0<p_1\le\ldots\le p_{\tau_o}\}$$

$$\text{OR} \quad \triangleq \{(0;p_i,\sigma_i), i\in\underset{\sim}{\nu}_o, 0<p_1<\ldots<p_{\nu_o}, \tau_o=\sum_{i=1}^{\nu_o}\tilde{\sigma}_i\} \tag{7.1c}$$

$$\mathcal{D}_\infty(F,G) \triangleq \{\hat{s}^{q_i}, 0<q_1\le q_2\le\ldots\le q_{\tau_\infty}\}$$

$$\text{OR} \quad \triangleq \{(\infty;q_i,\hat{\sigma}_i), i\in\underset{\sim}{\nu}_\infty, 0<q_1<\ldots<q_{\nu_\infty}, \tau_\infty=\sum_{i=1}^{\nu_\infty}\hat{\sigma}_i\} \tag{7.1d}$$

$$\mathcal{D}_\alpha(F,G) \triangleq \{(s-\alpha\hat{s})^{d_i}, \alpha\in\mathbb{C}-\{0\}, 0<d_1\le\ldots\le d_{\tau_\alpha}\}$$

$$\text{OR} \quad \triangleq \{(\alpha;d_i,\sigma_i), i\in\underset{\sim}{\nu}_\alpha, 0<d_1<\ldots<d_{\nu_\alpha}, \tau_\alpha=\sum_{i=1}^{\nu_\alpha}\sigma_i\} \tag{7.1e}$$

The second of the above descriptions will be referred to as the index set description of the corresponding invariant. For an $\alpha\in\mathbb{C}-\mathbb{R}$ we shall denote by $\mathcal{D}_{\alpha^*}(F,G)$ the complex conjugate set of $\mathcal{D}_\alpha(F,G)$ and by $\mathcal{D}_{\hat{\alpha}}(F,G)$ the inverse set of $\mathcal{D}_\alpha(F,G)$, i.e. the set of e.d. $(\hat{\alpha}s-\hat{s})^{d_i}$, $\hat{\alpha}=1/\alpha$ obtained from $\mathcal{D}_\alpha(F,G)$. The set of all distinct numbers $\Phi(F,G)=\{\alpha_i;\alpha_i\in\mathbb{C}\cup\{\infty\},\alpha_i\ne\alpha_j, i\in\underset{\sim}{\nu}\}$ for which $\mathrm{rank}\{\alpha_i F-G\}<\rho$ will be referred to as the root range of $sF-\hat{s}G$ and every α_i will be called a root representative. Clearly, $\Phi(F,G)$ is the set of distinct numbers (including infinity) which are associated with all possible e.d. of $sF-\hat{s}G$. The set $\bar{\Phi}(F,G)=\{(\alpha_i,\tau_{\alpha_i}),i\in\underset{\sim}{\upsilon}\}$, where τ_{α_i} is the algebraic multiplicity of $\alpha_i\in\Phi(F,G)$ in $\mathcal{D}_{\alpha_i}(F,G)$ will be referred to as the <u>root set</u> of (F,G). The set $\Phi(F,G)$ is clearly symmetric (i.e. if $\alpha_i\in\mathbb{C}-\mathbb{R}$ and $\alpha_i\in\Phi(F,G)$, then $\alpha_i^*\in\Phi(F,G)$) and the maximal subset of $\Phi(F,G)$ defined by $\Phi'(F,G)\triangleq$ $\triangleq\{\forall\alpha_i\in\Phi(F,G):\alpha_i\neq\alpha_i^*\}$ will be called the <u>half root range</u> of (F,G). Finally, $<\mathcal{D}(F,G)>\triangleq\cup\mathcal{D}_{\alpha_i}(F,G)$ for all $\alpha_i\in\Phi(F,G)$ denotes the set of all e.d. of $sF-\hat{s}G$.

For the pair (F,G), or with the pencil $sF-\hat{s}G$ we shall denote by $T_k(F,G)$, $\tilde{T}_k(F,G)$ the k-th order right-, left-Toeplitz matrices of (F,G) (see eqns. (5.6),(5.7)) and by N_r^k, N_ℓ^k the right-, left-null spaces of $T_k(F,G),\tilde{T}_k(F,G)$ correspondingly; if $\theta_k=\dim N_r^k$ and $\tilde{\theta}_k=\dim N_\ell^k$, then by $C_r(F,G)=\{\theta_k:k=-1,0,1,2,\dots,\theta_{-1}=\theta_o=0\}$ and $C_\ell(F,G)=\{\tilde{\theta}_k:k=-1,0,1,2,\dots,\tilde{\theta}_{-1}=\theta_o=0\}$ we shall denote the right-, left-(F,G) sequences of (F,G) respectively (see Chapter 5). The sequences $C_r(F,G),C_\ell(F,G)$ are Piecewise Arithmetic Progression sequences (PAPS) and their properties are characterised by Corollary (5.5), i.e.

$$\theta_k \le \frac{\theta_{k-1}+\theta_{k+1}}{2}\ ,\quad \underline{\theta}_{-1}=\theta_o=0\ ,\ \forall k=-1,0,1,\dots \tag{7.2}$$

The sequence $C_r(F,G)$ ($C_\ell(F,G)$) will be called <u>neutral</u>, if for $\forall k$, $\theta_k=0$, and shall be called <u>simple</u>, if its only singular point is k=0; clearly, if $C_r(F,G)$ ($C_\ell(F,G)$) is neutral, then $N_r\{sF-\hat{s}G\}=\{0\}$ ($N_\ell\{sF-\hat{s}G\}=\{0\}$) and if it simple, then $sF-\hat{s}G$ has only zero c.m.i. (zero r.m.i.). Because of the sign of the inequality (7.2), the sequences $C_r(F,G)$ and $C_\ell(F,G)$ will be referred to as <u>Nonincreasing</u> P.A.P.S. (NI.P.A.P.S.). For the sequence $C_r(F,G)$ ($C_\ell(F,G)$) we have the following properties.

<u>Proposition (7.1)</u>: Let $sF-\hat{s}G$ be a general pencil and let $\alpha\in\mathbb{C}$. Then, the following properties hold true:

$$C_r(F,G) = C_r(G,F) = C_r(F,G-\alpha F) = C_r(F-\alpha G,G) \tag{7.3}$$

Proof

By Theorem (5.2), it is clear that $C_r(F,G)$ is uniquely defined by $\mathcal{I}_c(F,G)$; thus, to prove the result we have to show the equivalent condition, i.e.

$$\mathcal{I}_c(F,G) = \mathcal{I}_c(G,F) = \mathcal{I}_c(F,G-\alpha F) = \mathcal{I}_c(F-\alpha G,G) \tag{7.4}$$

To prove (7.4), it is sufficient to show that the corresponding ordered pairs are $\mathcal{E}_{\mathcal{B}}$-equivalent (see Proposition (6.4)). It is obvious that $(F,G)\mathcal{E}_{\mathcal{B}}(G,F)$. If we define the projective transformation β:

$$\beta: \begin{bmatrix} s \\ \hat{s} \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda \\ \hat{\lambda} \end{bmatrix}, \ \alpha\in\mathbb{C} \tag{7.5a}$$

then for the induced transformation $b\in\mathcal{B}$ we have

$$b\circ(F,G) = (F,G-\alpha F) \tag{7.5b}$$

and thus $(F,G)\mathcal{E}_{\mathcal{B}}(F,G-\alpha F)$. Note that the invariance of $\mathcal{I}_c(F,G)$ under bilinear equivalence, is a general property under real, or complex projective transformations and this completes the proof. □

The above result provides the means to extend the notion of Piecewise Arithmetic Progression sequences, characterising the root representatives of a regular pencil (and thus the Segré characteristics) to the case of singular pencils.

Definition (7.1): Let $sF-\hat{s}G \in \mathbb{R}^{m\times n}[s,\hat{s}]$, $\operatorname{rank}_{\mathbb{R}(s,\hat{s})}\{sF-\hat{s}G\}=\rho\le\min\{m,n\}$. For $\forall\alpha$, $\hat{\alpha}\in\mathbb{C}$ we define:

(i) For $\forall$ k=1,2,... the matrices

$$P_{\alpha}^{1}(F,G) = G-\alpha F,\ldots,P_{\alpha}^{k}(F,G) = \begin{bmatrix} G-\alpha F & 0 & \cdots & 0 & 0 \\ -F & G-\alpha F & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & G-\alpha F & 0 \\ 0 & 0 & \cdots & -F & G-\alpha F \end{bmatrix} \in\mathbb{C}^{km\times kn} \tag{7.6a}$$

$$P^1(G,F) = F-\hat{\alpha}G,\ldots,P^k_{\hat{\alpha}}(G,F) = \begin{bmatrix} F-\hat{\alpha}G & 0 & \cdots & 0 & 0 \\ -G & F-\hat{\alpha}G & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & F-\hat{\alpha}G & 0 \\ 0 & 0 & \cdots & -G & F-\hat{\alpha}G \end{bmatrix} \tag{7.6b}$$

as the k-th order α-(F,G)-Toeplitz matrices (k-α-(F,G)-T.M.) and k-th order $\hat{\alpha}$-(G,F)-Toeplitz matrices (k-$\hat{\alpha}$-(G,F)-T.M.) of the pencil respectively. For $\hat{\alpha}=0$, we set $P^k_o(G,F)\triangleq P^k_\infty(F,G)$ and $P^k_\infty(F,G)$ shall be referred to as k-th order ∞-(F,G)-Toeplitz matrices (k-∞-(F,G)-T.M.) of the pencil.

(ii) Let $\mathcal{N}^k_{r,\alpha} \triangleq \mathcal{N}_r(P^k_\alpha(F,G))$, $\hat{\mathcal{N}}^k_{r,\hat{\alpha}} \triangleq \mathcal{N}_r(P^k_{\hat{\alpha}}(G,F))$, $\mathcal{N}^k_{\ell,\alpha} \triangleq \mathcal{N}_\ell(P^k_\alpha(F,G))$, $\hat{\mathcal{N}}^k_{\ell,\hat{\alpha}} \triangleq \mathcal{N}_\ell(P^k_{\hat{\alpha}}(G,F))$. The sequences defined by

$$J^r_\alpha(F,G) \triangleq \{n^\alpha_k : n^\alpha_{-1}=n^\alpha_o=0,\ n^\alpha_k=\dim \mathcal{N}^k_{r,\alpha} : k\geq 1\} \tag{7.7a}$$

$$J^r_{\hat{\alpha}}(G,F) \triangleq \{\hat{n}^{\hat{\alpha}}_k : \hat{n}^{\hat{\alpha}}_{-1}=\hat{n}^{\hat{\alpha}}_o=0,\ n^{\hat{\alpha}}_k=\dim \hat{\mathcal{N}}^k_{r,\hat{\alpha}} : k\geq 1\} \tag{7.7b}$$

will be referred to as the right-α-(F,G)-, right-$\hat{\alpha}$-(G,F)-sequences of the pencil respectively. For $\hat{\alpha}=0$, we set $J^r_o(G,F)\triangleq J^r_\infty(F,G)=\{n^\infty_k\}$ and $J^r_\infty(F,G)$ will be called the right-∞-(F,G)-sequence of the pencil. In a similar manner we may define the left-α-(F,G)-, left-$\hat{\alpha}$-(G,F)-sequences $J^\ell_\alpha(F,G), J^\ell_{\hat{\alpha}}(G,F)$, as well as $J^\ell_\infty(F,G)$ for the given pencil. □

Remark (7.1): If $sF-\hat{s}G$ is regular, then:

(i) For $\forall\alpha,\hat{\alpha}$ such that $\alpha,\hat{\alpha}\notin\Phi(F,G)$, the sequences $J^r_\alpha(F,G), J^r_{\hat{\alpha}}(G,F)$ are zero sequences (all their elements are zero).

(ii) For $\forall\alpha,\hat{\alpha}$ such that $\alpha,\hat{\alpha}\in\Phi(F,G)$, the sequences $J^r_\alpha(F,G), J^r_{\hat{\alpha}}(G,F)$ are nontrivial (they have at least one nonzero element); furthermore, $J^r_\alpha(F,G), J^r_{\hat{\alpha}}(G,F)$ are Piecewise Arithmetic Progression Sequences at $s=\alpha, s=\hat{\alpha}$ respectively. □

The sequences $J^r_\alpha(F,G), J^r_{\hat{\alpha}}(G,F)$ which are defined on a regular pencil $sF-\hat{s}G$ shall be denoted by $\bar{J}^r_\alpha(F,G), \bar{J}^r_{\hat{\alpha}}(G,F)$ and the piecewise arithmetic progression property is expressed by the relationship

$$n^\alpha_k \geq \frac{n^\alpha_{k-1}+n^\alpha_{k+1}}{2} \tag{7.8}$$

Because of the sign of the inequality (7.8), $\bar{J}^r_\alpha(F,G), \bar{J}^r_{\hat{\alpha}}(G,F)$ will be referred to as Nondecreasing P.A.P.S. (ND.P.A.P.S.). The sequences $J^r_\alpha(F,G)$, $J^r_{\hat{\alpha}}(G,F)$ play an important role in the study of the geometry of $S(F,G)$ systems and their properties are examined next.

Lemma (7.1): Let $sL_\varepsilon-\hat{s}\hat{L}_\varepsilon \in \mathbb{R}^{\varepsilon\times(\varepsilon+1)}[s,\hat{s}]$ be the pencil defined by $L_\varepsilon=[I_\varepsilon \,\vdots\, \underline{0}]$, $\hat{L}_\varepsilon=[\underline{0} \,\vdots\, I_\varepsilon]$, $\alpha,\hat{\alpha}\in\mathbb{C}$ and let $n^\alpha_k=\dim \mathcal{N}^k_{r,\alpha}$, $\hat{n}^{\hat{\alpha}}_k=\dim \hat{\mathcal{N}}^k_{r,\hat{\alpha}}$, $m^\alpha_k=\dim \mathcal{N}^k_{\ell,\alpha}$, $\hat{m}^{\hat{\alpha}}_k=\dim \hat{\mathcal{N}}^k_{\ell,\hat{\alpha}}$. For all $k=1,2,\ldots$, $n^\alpha_k=\hat{n}^{\hat{\alpha}}_k=k$ and $m^\alpha_k=\hat{m}^{\hat{\alpha}}_k=0$.

Proof

Consider the pencils $sF_1-\hat{s}G_1=s(L_\varepsilon-\alpha\hat{L}_\varepsilon)-\hat{s}\hat{L}_\varepsilon$ and $sF_1-\hat{s}G_2=sL_\varepsilon-\hat{s}(\hat{L}_\varepsilon-\alpha L_\varepsilon)$; clearly, $(F_1,G_1)E_{\mathcal{B}}(L_\varepsilon,\hat{L}_\varepsilon)$ and $(F_2,G_2)E_{\mathcal{B}}(L_\varepsilon,\hat{L}_\varepsilon)$. Since $sL_\varepsilon-\hat{s}\hat{L}_\varepsilon$ has as the only strict equivalence invariant the c.m.i. ε, then by Remark (6.19) we also have that $(F_1,G_1)E_{\mathcal{H}}(L_\varepsilon,\hat{L}_\varepsilon)$ and $(F_2,G_2)E_{\mathcal{H}}(L_\varepsilon,\hat{L}_\varepsilon)$. The investigation of the values of $n^\alpha_k, \hat{n}^\alpha_k$ is equivalent to a study of solutions of the matrix equations

$$F_i\underline{x}^i_1=0,\ G_i\underline{x}^i_1=F_i\underline{x}^i_2,\ldots,\ G_i\underline{x}^i_{k-1}=F_i\underline{x}^i_k,\ i=1,2$$

which because $(F_i,G_i)E_{\mathcal{H}}(L_\varepsilon,\hat{L}_\varepsilon)$ are equivalent to the set of equations (5.16). By Lemma (5.2), it is clear that for $\forall k$ the dimension of the solution space is k and thus $n^\alpha_k=\hat{n}^{\hat{\alpha}}_k=k$. The property $m^\alpha_k=\hat{m}^{\hat{\alpha}}_k=0$ follows from the linear independence of the rows of (F_i,G_i) for $i=1,2$. □

In the following we shall study the properties of the sequences $J^r_\alpha(F,G)$, $J^r_{\hat{\alpha}}(G,F), J^\ell_\alpha(F,G), J^\ell_{\hat{\alpha}}(G,F)$. The root range $\Phi(F,G)=\{\infty,0,\alpha_i\in\mathbb{C}-\{0\},\alpha_i\neq\alpha_j\}$ will be referred to the pencil $sF-G$, while $\phi(G,F)=\{0,\infty,\hat{\alpha}_i=\alpha_i^{-1}\in\mathbb{C}-\{0\},\hat{\alpha}_i\neq\hat{\alpha}_j\}$ will be referred to as the <u>inverse root range</u>, or as the root range of $F-\hat{s}G$ (the dual pencil). The results will be presented for the right sequences of the pencil sF-G (i.e. $J^r_\alpha(F,G), J^r_\infty(F,G)$), whereas the case of left sequences may be treated by invoking "transposed duality" and the case of $F-\hat{s}G$ by using "elementary divisor type duality".

<u>Proposition (7.2)</u>: Let $sF-G \in \mathbb{R}^{m\times n}[s]$, $p=\dim \mathcal{N}_r\{sF-G\}$, $\phi(F,G)$ be the root range of sF-G and let $n^\alpha_k=\dim \mathcal{N}^k_{r,\alpha}$, $\hat{n}^{\hat{\alpha}}_k=\dim \hat{\mathcal{N}}^k_{r,\hat{\alpha}}$ be the associated integers defined for some $\alpha,\hat{\alpha}\in\mathbb{C}\cup\{\infty\}$.

(a) For $\forall\alpha,\hat{\alpha}\in\mathbb{C}\cup\{\infty\}$, such that $\alpha\notin\phi(F,G)$, $\hat{\alpha}\notin\Phi(G,F)$, then $n^\alpha_k=\hat{n}^{\hat{\alpha}}_k=pk$.

(b) Let $\alpha\in\mathbb{C}$, $\alpha\in\Phi(F,G)$ and let $\mathcal{D}_\alpha(F,G)=\{(\alpha;d_i,\sigma_i), i\in\underset{\sim}{\nu}_\alpha, 0<d_1<\dots<d_{\nu_\alpha}, \tau_\alpha=\sum_{i=1}^{\nu_\alpha}\sigma_i\}$ be the e.d. set of sF-G associated with α. The dimensions n^α_k satisfy the following properties:

(i) If $d_i\le k<d_{i+1}$, then

$$n^\alpha_k = pk + \sum_{j=1}^{i} \sigma_j d_j + k \sum_{j=i+1}^{\nu_\alpha} \sigma_j$$

(ii) If $k<d_1$, then $n^\alpha_k=pk+k\sum_{j=1}^{\nu_\alpha}\sigma_j$.

(iii) If $k\ge d_{\nu_\alpha}$, then $n^\alpha_k=pk+\sum_{j=1}^{\nu\alpha}\sigma_j d_j$.

(c) Let $\infty\in\Phi(F,G)$ and let $\mathcal{D}_\infty(F,G)=\{(\infty;q_i,\hat{\sigma}_i), i\in\underset{\sim}{\nu}_\infty, 0<q_1<\dots<q_{\nu_\infty}, \tau_\infty=\sum_{i=1}^{\nu_\infty}\hat{\sigma}_i\}$ be the infinite e.d. set of sF-G. Then, $n^\infty_k=\hat{n}^o_k$ and $\hat{n}^o_k$ satisfy the following properties:

(i) If $q_i\le k<q_{i+1}$, then

$$n^\infty_k = \hat{n}^o_k = pk + \sum_{j=1}^{i} q_i\hat{\sigma}_i + k \sum_{j=i+1}^{\nu_\infty} \hat{\sigma}_j$$

(ii) If $k<q_1$, then $n^\infty_k=\hat{n}^o_k=pk+k\sum_{j=1}^{\nu_\infty}\hat{\sigma}_j$.

(iii) If $k\ge q_{\nu_\infty}$, then $n^\infty_k=\hat{n}^o_k=pk+\sum_{j=1}^{\nu_\infty}q_j\hat{\sigma}_j$.

<u>Proof</u>

It is clear that the numbers $n^\alpha_k, \hat{n}^{\hat{\alpha}}_k$ are invariant under strict equivalence transformations applied on sF-G; thus, the properties of $n_k^{\alpha},\hat{n}_k^{\hat{\alpha}}$ may be studied on any element of the orbit $\mathcal{E}_{\mathcal{H}}(F,G)$. Let (R,Q) be a pair of $\mathcal{E}_{\mathcal{H}}$ transformations which reduce sF-G to

$$sF'-G' = R(sF-G)Q = \text{block-diag}\{sF_c-G_c; sF_r-G_r; s\bar{F}-\bar{G}\} \tag{7.9a}$$

where $s\bar{F}-\bar{G}, sF_c-G_c, sF_r-G_r$ are regular-, right-, left-restrictions correspondingly of sF-G. By the block-diagonal decomposition of sF'-G' it follows that the null spaces of $P_{\alpha}^{k}(F',G'), P_{\hat{\alpha}}^{k}(G',F')$ may be expressed as direct sums of the corresponding null spaces of the matrices defined on the subpencils $sF_c-G_c, sF_r-G_r, s\bar{F}-\bar{G}$; thus, if $n_{\alpha}^{k}, n'^{k}_{\alpha}, n^{k}_{c,\alpha}, n^{k}_{r,\alpha}, n''^{k}_{\alpha}$ and $\hat{n}^{k}_{\hat{\alpha}}, \hat{n}'^{k}_{\hat{\alpha}}, \hat{n}^{k}_{c,\hat{\alpha}}, \hat{n}^{k}_{r,\hat{\alpha}}, \hat{n}''^{k}_{\hat{\alpha}}$ are the numbers defined for α and $\hat{\alpha}\in\mathbb{C}$ on the pencils $sF-G, sF'-G', sF_c-G_c, sF_r-G_r, s\bar{F}-\bar{G}$ correspondingly then

$$n_{\alpha}^{k} = n'^{k}_{\alpha} = n^{k}_{c,\alpha} + n^{k}_{r,\alpha} + n''^{k}_{\alpha} \quad \forall k \tag{7.9b}$$

$$\hat{n}^{k}_{\hat{\alpha}} = \hat{n}'^{k}_{\hat{\alpha}} = \hat{n}^{k}_{c,\hat{\alpha}} + \hat{n}^{k}_{r,\hat{\alpha}} + \hat{n}''^{k}_{\hat{\alpha}} \quad \forall k \tag{7.9c}$$

Note that sF_c-G_c, sF_r-G_r may always be considered in the Kronecker form and thus by Lemma (7.1) it is readily shown that

$$n^{k}_{c,\alpha} = \hat{n}^{k}_{c,\hat{\alpha}} = pk, \quad n^{k}_{r,\alpha} = \hat{n}^{k}_{r,\hat{\alpha}} = 0 \quad \forall k \tag{7.9d}$$

The properties of the numbers $n''^{k}_{\alpha}, \hat{n}''^{k}_{\hat{\alpha}}$ are described by Corollary (4.4) and the result is readily established. □

Remark (7.2): If $n^{k}_{\alpha}, \hat{n}^{k}_{\hat{\alpha}}$ and $n''^{k}_{\alpha}, \hat{n}''^{k}_{\hat{\alpha}}$ are the numbers defined on the pencils sF-G and $s\bar{F}-\bar{G}$, where $s\bar{F}-\bar{G}$ is a regular restriction of sF-G, then

$$n^{k}_{\alpha} = pk + n''^{k}_{\alpha} \quad \forall\alpha\in\mathbb{C},\ \forall k=1,2,\ldots \tag{7.10a}$$

$$\hat{n}^{k}_{\hat{\alpha}} = pk + \hat{n}''^{k}_{\hat{\alpha}} \quad \forall\hat{\alpha}\in\mathbb{C},\ \forall k=1,2,\ldots \tag{7.10b}$$

where $p=\dim \mathcal{N}_r\{sF-G\}$. □

The above remark allows the extension of the results derived for the Piecewise Arithmetic Progression Sequences of a regular pair (F,G) at $s=\alpha$, to the case of singular pairs.

a <u>Proposition (7.3)</u>: The sequences $\mathcal{J}_{\alpha}^{r}(F,G), \mathcal{J}_{\hat{\alpha}}^{r}(G,F)$ defined for some $\alpha\in\Phi(F,G), \hat{\alpha}\in\Phi(G,F), \alpha, \hat{\alpha}\in\mathbb{C}$ on a general pencil sF-G are ND.P.A.P.S. and for all k=1,2,... satisfy the properties:

$$n_{\alpha}^{k} \ge \frac{n_{\alpha}^{k-1}+n_{\alpha}^{k+1}}{2}, \quad \hat{n}_{\hat{\alpha}}^{k} \ge \frac{\hat{n}_{\hat{\alpha}}^{k-1}+\hat{n}_{\hat{\alpha}}^{k+1}}{2} \tag{7.11}$$

In particular, we have that:

(i) Strict inequality holds if and only if k is the degree of an e.d. of $sF-G, F-\hat{s}G$ at $s=\alpha, \hat{s}=\hat{\alpha}$ correspondingly.

(ii) Inequality holds if and only if k is not the degree of an e.d. of $sF-G, F-\hat{s}G$ at $s=\alpha, \hat{s}=\hat{\alpha}$ correspondingly.

<u>Proof</u>

By noting that $n_{\alpha}^{k}-n_{\alpha}^{k-1}=p+n_{\alpha}^{\prime\prime k}-n_{\alpha}^{\prime\prime k-1}, n_{\alpha}^{\prime\prime -1}=0$ and by using Corollary (4.5) the result follows.

□

If $s\bar{F}-\bar{G}$ denotes any regular restriction of sF-G, $I_p=\{n_k : n_{-1}=n_o=0, n_k=pk, k\ge 1\}$, then Remark (7.2) implies that

$$\mathcal{J}_{\alpha}^{r}(F,G) = I_p + \mathcal{J}_{\alpha}^{r}(\bar{F},\bar{G}) \quad \forall\alpha\in\mathbb{C} \tag{7.12a}$$

$$\mathcal{J}_{\hat{\alpha}}^{r}(G,F) = I_p + \mathcal{J}_{\hat{\alpha}}^{r}(\bar{G},\bar{F}) \quad \forall\hat{\alpha}\in\mathbb{C} \tag{7.12b}$$

Clearly, $\mathcal{J}_{\alpha}^{r}(\bar{F},\bar{G}), \mathcal{J}_{\hat{\alpha}}^{r}(\bar{G},\bar{F})$ are neutral (all elements zero) if $\alpha\notin\phi(F,G)$, $\hat{\alpha}\notin\phi(G,F)$; otherwise they are nontrivial and their properties (singularities and corresponding gaps) characterise the e.d. structure (Segre characteristic) of $sF-G, F-\hat{s}G$ at the corresponding root representative. To define the sequences $\mathcal{J}_{\alpha}^{r}(\bar{F},\bar{G}), \mathcal{J}_{\hat{\alpha}}^{r}(\bar{G},\bar{F})$ from the sequences $\mathcal{J}_{\alpha}^{r}(F,G), \mathcal{J}_{\hat{\alpha}}^{r}(G,F)$ correspondingly we do not have to work out a regular restriction of sF-G but merely the number p which uniquely defines the sequence I_p. The number $p=\dim N_r\{sF-G\}$ may be readily defined by the following results.

Proposition (7.4): Let sF-G be a general pencil, $p=\dim N_r\{sF-G\}$ and $C_r(F,G)$ the corresponding sequence associated with the pencil. The number p may be defined by either of the following two methods:

(i) Let τ be an integer such that for $\forall k\geq\tau$, $\theta_k=(\theta_{k+1}+\theta_{k-1})/2$. Then $p=\theta_k-\theta_{k-1}$.

(ii) Let $\alpha\in\mathbb{C}$ be a number such that $\alpha\notin\Phi(F,G)$, then $p=\dim N_r\{G-\alpha F\}=$ $=\dim N_r\{F-\alpha G\}$.

Proof

Part (i) follows from Remark (5.7), whereas part (ii) follows by inspection of the Kronecker form. □

Clearly by selecting a few random numbers $\alpha_i\in\mathbb{C}$ and computing $\min\{\dim N_r\{G-\alpha_i F\}\}$ we may compute p. Then the sequences $J^r_\alpha(\bar{F},\bar{G}),J^r_{\hat{\alpha}}(\bar{G},\bar{F})$ may be found by

$$J^r_\alpha(\bar{F},\bar{G}) = J^r_\alpha(F,G) - I_p \;,\; \forall\alpha\in\mathbb{C} \tag{7.13a}$$

$$J^r_{\hat{\alpha}}(\bar{G},\bar{F}) = J^r_{\hat{\alpha}}(G,F) - I_p \;,\; \forall\alpha\in\mathbb{C} \tag{7.13b}$$

If $J^r_\alpha(\bar{F},\bar{G}),J^r_{\hat{\alpha}}(\bar{G},\bar{F})$ are non-neutral, then the procedures described in Chapter (4) may be used for the computation of the Segre characteristic for $\forall\alpha\in\Phi(F,G)$. A combination of the procedures described in Chapter (4) and Chapter (5) for the computation of the Weierstrass canonical form of a regular pencil and for the computation of the sets $I_c(F,G),I_r(F,G)$ of a singular pencil correspondingly readily yields a procedure for computing the Kronecker form of a pencil without resorting to the use of transformations. A summary of this procedure is given below.

The Piecewise Arithmetic Progression Sequences Approach for the computation of the Kronecker canonical form

Let $sF-G \in \mathbb{R}^{m\times n}[s]$ be a general pencil. The Kronecker form of sF-G may be computed without using strict equivalence transformations as follows:

Step (1): Compute the sequences $\mathcal{C}_r(F,G), \mathcal{C}_\ell(F,G)$ associated with the pencil and by use of the PAPSD of Chapter (5) compute the sets $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$, as well as the numbers $p=\dim \mathcal{N}_r\{sF-G\}$ and $t=\dim \mathcal{N}_\ell\{sF-G\}$. Then the $\text{rank}_{\mathbb{R}(s)}\{sF-G\}=\rho$ is defined by $\rho=m-t=m-n+p$ if $m\leq n$ and $\rho=n-m+t=n-p$ if $m\geq n$.

Step (2): The root range $\Phi(F,G)$ is computed as follows:

(a) If $\dim \mathcal{N}_r(F)>p$, then $\infty\in\Phi(F,G)$; otherwise $\infty\notin\Phi(F,G)$.

(b) If $\dim \mathcal{N}_r(G)>p$, then $0\in\Phi(F,G)$; otherwise $0\notin\Phi(F,G)$.

(c) Let $sF^\rho-G^\rho$ denote the $\rho\times\rho$ subpencils of sF-G, which correspond to an $\{a\}$ ρ-set of columns and a $\{b\}$ ρ-set of columns of sF-G. There exists at least one subpencil $sF^\rho-G^\rho$ for which $\mathcal{C}_r(F^\rho,G^\rho)$ is neutral (and thus $\mathcal{C}_\ell(F^\rho\ G^\rho)$ is also neutral); such subpencils will be referred to as <u>ρ-regular subpencils</u>. If $sF^\rho-G^\rho$ is a ρ-regular subpencil find $\det(sF^\rho-G^\rho)=a(s)$ and let $\sigma(F^\rho,G^\rho)$ be the distinct roots of a(s). For $\forall\beta\in\sigma(F^\rho,G^\rho)$ compute $G-\beta F$. If $\dim \mathcal{N}_r(G-\beta F)>\rho$, then $\beta\in\Phi(F,G)$, otherwise $\beta\notin\Phi(F,G)$. The procedure yields the set $\Phi(F,G)$.

Step (3): For $\forall\alpha\in\Phi(F,G), \alpha\in\mathbb{C}$, compute the sequence $\mathcal{J}^r_\alpha(F,G)$ and by (7.13a) the sequence $\mathcal{J}^r_\alpha(\bar{F},\bar{G})$. Similarly for $\alpha=\infty\in\Phi(F,G)$ compute the sequence $\mathcal{J}^r_o(G,F)$ and thus by (7.13b) the sequence $\mathcal{J}^r_o(\bar{G},\bar{F})$. From the sequences $\mathcal{J}^r_\alpha(\bar{F},\bar{G}), \mathcal{J}^r_o(\bar{G},\bar{F})$ compute the Segré characteristics for $\alpha\in\mathbb{C}$, or $\alpha=\infty$ by using either of the methods described in Chapter (4). The procedure yields the sets $\mathcal{D}_\infty(F,G), \mathcal{D}_\alpha(F,G)$ of the pencil which together with the sets $\mathcal{I}_c(F,G)$, $\mathcal{I}_r(F,G)$ define the Kronecker form. □

The sequences $\mathcal{C}_r(F,G), \mathcal{C}_\ell(F,G)$ and $\mathcal{J}^r_\alpha(F,G)$, for $\forall\alpha\in\Phi(F,G)$ uniquely characterise the strict equivalence orbit $\mathcal{E}_{\mathcal{H}}(F,G)$; this characterisation is summarised by the following result.

Theorem (7.1): Let $sF-G \in \mathbb{R}^{m\times n}[s]$ be a general pencil and let $\Phi(F,G)= =\{\alpha_i, \alpha_i \in \mathbb{C}\cup\{\infty\}, i\in \underset{\sim}{\nu}\}$ be the root range of sF-G. The set $\Phi(F,G)$ and the sequences $\mathcal{J}^r_\alpha(F,G), \forall \alpha\in\Phi(F,G), \mathcal{C}_r(F,G), \mathcal{C}_\ell(F,G)$ form a complete set of invariants for the strict equivalence class $\mathcal{E}_{\mathcal{H}}(F,G)$. □

By interpreting the strict equivalence invariants of $\mathcal{E}_{\mathcal{H}}(F,G)$ in terms of the corresponding sequences a unification of the different notions of invariants is achieved, since each one of them is characterised by the properties of a Piecewise Arithmetic Progression Sequence. The Kronecker form may then be interpreted as a canonical form describing the singular points and associated gaps of the corresponding PAPS.

7.3 The geometry of the subspaces of the domain of (F,G)

7.3.1 Introduction: Background notation and definitions

The algebraic, number theoretic and frequency-space relativistic properties of matrix pencils have a natural geometric interpretation; it is the aim of the present section to study these properties and to demonstrate that the geometric properties may be naturally derived as a byproduct of the underlying algebraic and number theoretic aspects. The effect of frequency-space transformations on the geometric properties will be examined in the following chapter. Central to our study is the classification of geometric properties of the subspaces of the domain of the pair (F,G).

Let $\mathcal{U}$ and $\mathcal{W}$ be vector spaces over a common field $\mathcal{F}$ (in our present study $\mathcal{F}=\mathbb{R}$) and let $\mathcal{L}(\mathcal{U};\mathcal{W})$ denote the set of all linear mappings from $\mathcal{U}$ into $\mathcal{W}$. The set $\mathcal{L}(\mathcal{U};\mathcal{W})$ is a linear vector space under addition and scalar multiplication. If $(\mathscr{f},\mathscr{g})\in L(\mathcal{U};\mathcal{W})\times L(\mathcal{U};\mathcal{W})$, then the pencil of linear operators $s\mathscr{f}-\hat{s}\mathscr{g}$ defined on the pair $(\mathscr{f},\mathscr{g})$ is a mapping of $\mathcal{U}$ into $\mathcal{W}$. For a definite choice of bases $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$ in these spaces the pencil of operators $s\mathscr{f}-\hat{s}\mathscr{g}$ corresponds to a pencil of rectangular matrices $sF-\hat{s}G$ (of dimension $m\times n$, $m=\dim\mathcal{W}$, $n=\dim\mathcal{U}$); the representation is illustrated by the following commutative diagram

$$\begin{tikzcd}
\begin{matrix}\mathcal{V}\\ \mathcal{B}_{\mathcal{V}}\end{matrix} \arrow[r, "s\mathscr{f}-\hat{s}\mathscr{g}"] \arrow[d, "\mathscr{r}_{u}"] & \begin{matrix}\mathcal{W}\\ \mathcal{B}_{\mathcal{W}}\end{matrix} \arrow[d, "\mathscr{r}_{\mathcal{W}}"] \\
F^{n} \arrow[r, "sF-\hat{s}G"] & F^{m}
\end{tikzcd}$$

Figure (7.1)

where $\mathscr{r}_{\mathcal{U}},\mathscr{r}_{\mathcal{W}}$ are the representation maps of $\mathcal{U},\mathcal{W}$ with respect to the given bases $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$. Thus $sF-\hat{s}G$ is a matrix representation of $s\mathscr{f}-\hat{s}\mathscr{g}$ with respect to $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$ and under a change of bases in $\mathcal{U},\mathcal{W}$, the new matrix representation of $s\mathscr{f}-\hat{s}\mathscr{g}$ becomes $R(sF-\hat{s}G)Q$ where R,Q are square nonsingular matrices expressing the coordinate transformations in the domain and codomain of $s\mathscr{f}-\hat{s}\mathscr{g}$. Clearly, strict equivalence of matrix pencils is equivalent to a study of the pencil of linear operators $s\mathscr{f}-\hat{s}\mathscr{g}$ and the set of strict equivalence invariants characterise uniquely the pencil of operators. It is therefore expected that the geometric properties of the subspaces of $\mathcal{U}$ under the $(\mathscr{f},\mathscr{g})$-pair mapping to be intimately related to the set of strict equivalence invariants which characterise $s\mathit{f}-\hat{s}\mathit{g}$. In the following it will be assumed that $\mathcal{U}=\mathbb{R}^n$, $\mathcal{W}=\mathbb{R}^m$ and thus $sF-\hat{s}G$ is the pencil of interest which is associated with the ordered pair $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$. If $\mathcal{V}$ is a subspace of $\mathbb{R}^n$ and $(F\mathcal{V},G\mathcal{V})$ denotes the ordered pair of the images of $\mathcal{V}$ under (F,G), then the fundamental geometric questions which have to be examined are those revolving around the relationships between the subspaces $F\mathcal{V}$,$G\mathcal{V}$ represented in Figure (7.2).

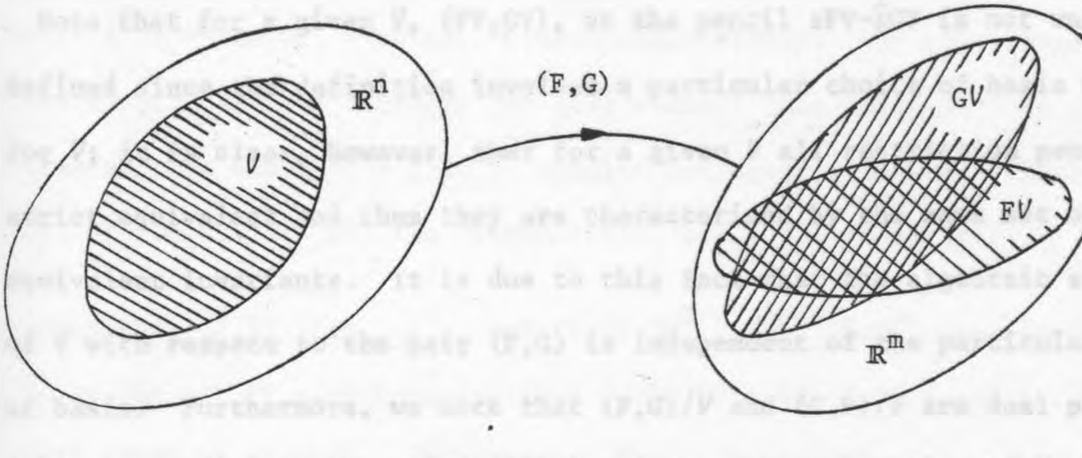

Figure (7.2)

The key tool in our study is the notion of the $\mathcal{V}$-restricted ordered pair, or alternatively of the $\mathcal{V}$-restricted matrix pencil. The notion of the $\mathcal{V}$-restricted ordered pair of maps is an extension of the standard notion defined on linear maps. Thus, let $\mathit{f}:\mathcal{U}\to\mathcal{W}$ be a linear map, $\mathcal{V}\subset\mathcal{U}$ be a subspace with insertion map $u:\mathcal{V}\to\mathcal{U}$. Let $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$ be a basis for $\mathcal{U}$,$\mathcal{W}$ respectively and let F and V be the matrix representations of f and u correspondingly with respect to $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$. The <u>restriction of f to $\mathcal{V}$</u> is the map $\mathit{f}/\mathcal{V}:\mathcal{V}\to\mathcal{W}$ defined by $\mathit{f}/\mathcal{V}\overset{\Delta}{=}\mathit{f}u$ and the matrix representation of $\mathit{f}/\mathcal{V}$ with respect to the bases $\mathcal{B}_{\mathcal{U}},\mathcal{B}_{\mathcal{W}}$ is defined by

$$[\mathit{f}/\mathcal{V}] = [\mathit{f}][v] = FV \tag{7.14}$$

where by $[\cdot]$ denote the operation of matrix representation. This notion may be readily extended to ordered pairs (F,G) (or (f,g)) and thus to matrix pencils as follows:

<u>Definition (7.2)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace of $\mathbb{R}^n$ and V a basis matrix of $\mathcal{V}$ (relative to the standard basis of $\mathbb{R}^n$). The pair (FV,GV) will be called a <u>$\mathcal{V}$-restricted ordered pair</u>; the associated pencils sFV-GV,FV-$\hat{s}$GV will be termed <u>(F,G)-, (G,F)-$\mathcal{V}$ restriction pencils</u> and shall be denoted by $(F,G)/\mathcal{V},(G,F)/\mathcal{V}$ respectively. □

Note that for a given $\mathcal{V}$, (FV,GV), or the pencil sFV-$\hat{s}$GV is not uniquely defined since the definition involves a particular choice of basis matrix for $\mathcal{V}$; it is clear, however, that for a given $\mathcal{V}$ all restriction pencils are strict equivalent and thus they are characterised by the same set of strict equivalent invariants. It is due to this fact that the algebraic structure of $\mathcal{V}$ with respect to the pair (F,G) is independent of the particular choice of basis. Furthermore, we note that $(F,G)/\mathcal{V}$ and $(G,F)/\mathcal{V}$ are dual pencils (elementary divisor type of duality) and thus the properties of $\mathcal{V}$ with respect to (F,G) may be studied either in terms of sFV-GV, or in terms of FV-$\hat{s}$GV. The uniqueness of the characterisation of $\mathcal{V}$ with respect to (F,G), in terms of the strict equivalence of $(F,G)/\mathcal{V}$, or $(G,F)/\mathcal{V}$ has led to the algebraic characterisation of $\mathcal{V}$ [Kar. 1]. Before we introduce this characterisation we give some useful notation.

The sets of invariants defined by eqns.(7.1) will be denoted by $I_c(F,G)$, $I_r(F,G),\mathcal{D}_o(F,G),\mathcal{D}_\infty(F,G),\mathcal{D}_\alpha(F,G)$, when they are referred to the (F,G) pair, or the pencil sF-G ($\hat{s}$=1) and by $I_c(G,F),I_r(G,F),\mathcal{D}_\infty(G,F),\mathcal{D}_o(G,F),\mathcal{D}_{\hat{\alpha}}(G,F)$, when they are referred to the pair (G,F), or the pencil F-$\hat{s}$G (s=1). Whenever there is no ambiguity about the pair (F,G), or (G,F), the types of invariants of sF-G will be denoted in short by $I_c,I_r,\mathcal{D}^s_o,\mathcal{D}^s_\infty,\mathcal{D}^s_\alpha$ and those of F-$\hat{s}$G by $I_c,I_r,\mathcal{D}^{\hat{s}}_\infty,\mathcal{D}^{\hat{s}}_o,\mathcal{D}^{\hat{s}}_\alpha$ $(\hat{\alpha}=\alpha^{-1})$. A set of the above type (i.e. I_c,I_r, $\mathcal{D}^s_o,\mathcal{D}^s_\infty,\mathcal{D}^s_\alpha$) will be called <u>prime</u>, if it contains only one element (i.e. one c.m.i., or r.m.i., or one e.d. respectively). The root range of a pair (F,G) will be denoted in short by Φ, whenever there is no ambiguity about the pair; the inverse root range of (F,G), or the root range of (G,F) shall be denoted by $\hat{\Phi}$. The root range Φ, or the corresponding pair (F,G) shall be called <u>∞-proper</u>, <u>o-proper</u> and shall be denoted by Φ_∞, Φ_o, respectively, if $\infty \notin \Phi, o \notin \Phi$ correspondingly; if Φ is both ∞-proper and 0-proper, then it will be referred to as <u>proper</u> and shall be denoted by Φ_p. Clearly, by the e.d. type of duality, if Φ is o-proper, ∞-proper, proper, then $\hat{\Phi}$ is ∞-proper, o-proper, proper respectively. The root range will be called <u>simple</u>, if Φ contains one real element (the point at is also treated this way), or a pair of complex conjugate elements. Finally, a set of r.m.i. with all its elements zero will be denoted by $\mathcal{I}_r^o$.

<u>Definition (7.3) [Kar. -1]</u>: Let $(F,G) \in \mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace, $\mathcal{I}_{\mathcal{V}}$ be the set of strict equivalence invaraints of $(F,G)/\mathcal{V}$ and let Φ be the root range of $(F,G)/\mathcal{V}$. The subspace $\mathcal{V}$ will be called:

(a) <u>Φ-(F,G)-elementary divisor subspace</u> (Φ-(F,G)-e.d.s.), if $\mathcal{I}_{\mathcal{V}}=\{\mathcal{D}^s_{\alpha_i}, \forall\alpha_i\in\Phi$ and possibly $\mathcal{I}_r^o\}$. If Φ is o-proper, ∞-proper, proper, then $\mathcal{V}$ will be called respectively <u>o-proper</u>, <u>∞-proper</u>, <u>proper</u>, <u>Φ-(F,G)-e.d.s.</u> and shall be denoted by Φ_o,-(F,G)-e.d.s., Φ_∞,-(F,G)-e.d.s., Φ_p,-(F,G)-e.d.s. correspondingly. If the set Φ is simple, then the corresponding subspaces subspaces will be referred to as <u>simple</u>; if $\Phi=\{0\},\{\infty\},\{\alpha,\alpha\in\mathbb{R}-\{0\}\}$ or, $\{\alpha,\alpha^*,\alpha,\alpha^*\in\mathbb{C}-\{0\}\}$, then the corresponding subspaces will be denoted by 0-(F,G)-e.d.s., ∞-(F,G)-e.d.s., α-(F,G)-e.d.s., (α,α^*)-(F,G)-e.d.s. respectively. If Φ is simple and the corresponding set of e.d. is prime, then $\mathcal{V}$ will be called <u>prime ϕ-(F,G)-e.d.s.</u>

(b) <u>$\mathcal{I}_c$-(F,G)-column minimal index subspace</u> ($\mathcal{I}_c$-(F,G)-c.m.i.s.) if $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_c$ and possibly $\mathcal{I}_r^o\}$. If $\mathcal{I}_c$ is prime and $\mathcal{I}_c=\{\varepsilon\}$, then $\mathcal{V}$ will be called <u>prime</u> and shall be denoted by ε-(F,G)-c.m.i.s.

(c) *I_r-(F,G)-row minimal index subspace* (I_r-(F,G)-c.m.i.s.), if $I_V=\{I_r\}$. If I_r contains at least one nonzero element, then it will be called *nonreduced*, otherwise ($I_r=I_r^o$) it will be referred to as *reduced*. □

By the e.d. type duality we have the following obvious result.

Proposition (7.5): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $V\subset\mathbb{R}^n$ be a subspace (F,G)/V, (G,F)/V be the associated restriction pencils and let $\Phi,\hat{\Phi}$ be the corresponding root ranges of the two pencils. The pencils (F,G)/V,(G,F)/V are dual, $\hat{\Phi}$ is the inverse root range of Φ and the following properties hold true:

(i) V is Φ_o,-(F,G)-e.d.s., iff it is $\hat{\Phi}_\infty$,-(G,F)-e.d.s.

(ii) V is Φ_∞,-(F,G)-e.d.s., iff it is $\hat{\Phi}_o$,-(G,F)-e.d.s.

(iii) V is Φ_p,-(F,G)-e.d.s., iff it is $\hat{\Phi}_p$,-(G,F)-e.d.s.

(iv) V is I_c,-(F,G)-c.m.i.s., iff it is I_c,-(G,F)-c.m.i.s.

(v) V is I_r,-(F,G)-r.m.i.s., iff it is I_c,-(G,F)-r.m.i.s. □

Remark (7.3): If V is simple Φ-(F,G)-e.d.s. and the corresponding e.d. set is $\{\mathcal{D}^s_o\},\{\mathcal{D}^s_\infty\},\{\mathcal{D}^s_\alpha,\alpha\in\mathbb{R}-\{0\}\}$, or, $\{\mathcal{D}^s_\alpha,\mathcal{D}^s_{\alpha*},\alpha,\alpha*\in\mathbb{C}-\{0\}\}$, then V is also simple $\hat{\Phi}$-(G,F)-e.d.s. with an e.d. set $\{\hat{\mathcal{D}}^s_\infty\},\{\hat{\mathcal{D}}^s_o\},\{\mathcal{D}^s_{\hat{\alpha}},\hat{\alpha}=\alpha^{-1}\}$, or, $\{\mathcal{D}^s_{\hat{\alpha}},\mathcal{D}^s_{\hat{\alpha}*},\hat{\alpha}=\alpha^{-1},\hat{\alpha}*=\alpha*^{-1}\}$ respectively. □

Any subspace $V\subset\mathbb{R}^n$ is characterised by the set I_V which is associated with (F,G)/V. The block diagonal decomposition of the Kronecker canonical form of (F,G)/V, clearly suggests that V may be decomposed into a direct sum of prime o-(F,G)-e.d.s., ∞-(F,G)-e.d.s., α-(F,G)-e.d.s., $\alpha\in\mathbb{R}-\{0\}$, $(\alpha,\alpha*)$-(F,G)-e.d.s., $\alpha,\alpha*\in\mathbb{C}-\{0\}$, ε-(F,G)-c.m.i.s. and nonreduced ζ-(F,G)-r.m.i.s. Subspaces of the above type will be referred to in short as *prime invariant subspaces* of the pair (F,G) and their study is of crucial importance in our attempt to classify the subspaces of the domain of (F,G).

7.3.2 The structure of simple invariant subspaces of (F,G)

Our aim in this section is to derive a set of geometric and number theoretic conditions which characterise the prime (simple) invariant subspaces of (F,G). The results derived here provide the means for the classification of the subspaces $\mathcal{V}$ of the domain of (F,G).

Theorem (7.2): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let dim $\mathcal{V}$=d. Necessary and sufficient conditions for $\mathcal{V}$ to be a prime α-(F,G)-e.d.s., $\alpha\in\mathbb{R}-\{0\}$ and thus a prime $\hat{\alpha}$-(G,F)-e.d.s., $\hat{\alpha}=\alpha^{-1}$ are:

(i) $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $\mathcal{N}_r(G)\cap\mathcal{V}=0$.

(ii) There exist bases $\beta_{\alpha}=\{\underline{x}_i, i\in\underset{\sim}{d}\}$, $\beta_{\alpha}=\{\hat{\underline{x}}_i, i\in\underset{\sim}{d}\}$ and $\alpha\in\mathbb{R}-\{0\}$ such that

$$G\underline{x}_i = \alpha F\underline{x}_i + F\underline{x}_{i-1},\ i\in\underset{\sim}{d},\ \underline{x}_o=\underline{0} \tag{7.15a}$$

$$F\hat{\underline{x}}_i = \hat{\alpha}G\hat{\underline{x}}_i + G\hat{\underline{x}}_{i-1},\ i\in\underset{\sim}{d},\ \hat{\underline{x}}_o=\underline{0},\ \hat{\alpha}=\hat{\alpha}^{-1} \tag{7.15b}$$

Proof

If V is a basis matrix for $\mathcal{V}$, then the pair (FV,GV) is characterised by an e.d. $(s-\alpha)^d$ and possibly I_r^o. By inspection of the Kronecker form it follows that $\mathcal{N}_r(FV)=\mathcal{N}_r(GV)=0$ and thus $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $\mathcal{N}_r(G)\cap\mathcal{V}=0$. By Proposition (4.1), the existence of the bases $\{\underline{x}_i, i\in\underset{\sim}{d}\}$ and $\{\hat{\underline{x}}_i, i\in\underset{\sim}{d}\}$ for which conditions (7.15a),(7.15b) are satisfied is established; this proves the necessity. The sufficiency is established as follows: Conditions (i) imply that for every basis matrix V of $\mathcal{V}$, $\mathcal{N}_r(FV)=\mathcal{N}_r(GV)=0$ and thus the pencil (F,G)/$\mathcal{V}$ has no c.m.i. and no zero, infinite e.d. By Proposition (4.2) the existence of the basis $\{\underline{x}_i, i\in\underset{\sim}{d}\}$,$\{\hat{\underline{x}}_i, i\in\underset{\sim}{d}\}$ satisfying the chain conditions (7.15a),(7.15b) respectively imply the existence of an e.d. $(s-\alpha)^d$ for (F,G)/$\mathcal{V}$, or $(\hat{s}-\hat{\alpha})^d$ for (G,F)/$\mathcal{V}$. By inspection of the dimensions of (F,G)/$\mathcal{V}$, it follows that if there are r.m.i., then all of them must be zero.

□

Prime (α,α^*)-(F,G)-e.d.s. may be characterised by a similar result as it is stated below. Note, that since $\alpha=\sigma+j\omega, \alpha^*=\sigma-j\omega\in\mathbb{C}$, such subspaces may also be denoted as (σ,ω)-(F,G)-e.d.s.

Corollary (7.1): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $V\subset\mathbb{R}^n$ be a subspace and let $\dim V=2d$. Necessary and sufficient conditions for V to be a prime (σ,ω)-(F,G)-e.d.s. and thus a $(\hat{\sigma},\hat{\omega})$-$(G,F)$-e.d.s., where $\hat{\sigma}+j\hat{\omega}=(\sigma+j\omega)^{-1}$ are:

(i) $N_r(F)\cap V=0$ and $N_r(G)\cap V=0$.

(ii) There exist bases $B_{\sigma,\omega}=\{(\underline{a}_i,\underline{b}_i), i\in\underset{\sim}{d}\}$, $B_{\hat{\sigma},\hat{\omega}}=\{(\hat{\underline{a}}_i,\hat{\underline{b}}_i), i\in\underset{\sim}{d}\}$ and $\alpha=\sigma+j\omega\in\mathbb{C}-\mathbb{R}$, $\hat{\alpha}=\alpha^{-1}=\hat{\sigma}+j\hat{\omega}$ such that

$$\begin{cases} G\underline{a}_i = F\{\sigma\underline{a}_i-\omega\underline{b}_i\} + F\underline{a}_{i-1}, \ i\in\underset{\sim}{d}, \ \underline{a}_o=\underline{0} \\ G\underline{b}_i = F\{\omega\underline{a}_i+\sigma\underline{b}_i\} + F\underline{b}_{i-1}, \ i\in\underset{\sim}{d}, \ \underline{b}_o=\underline{0} \end{cases} \tag{7.16a}$$

$$\begin{cases} F\hat{\underline{a}}_i = G\{\hat{\sigma}\hat{\underline{a}}_i-\hat{\omega}\hat{\underline{b}}_i\} + G\hat{\underline{a}}_{i-1}, \ i\in\underset{\sim}{d}, \ \hat{\underline{a}}_o=\underline{0} \\ F\hat{\underline{b}}_i = G\{\hat{\omega}\hat{\underline{a}}_i+\hat{\sigma}\hat{\underline{b}}_i\} + G\hat{\underline{b}}_{i-1}, \ i\in\underset{\sim}{d}, \ \hat{\underline{b}}_o=\underline{0} \end{cases} \tag{7.16b}$$

□

The proof of the above Corollary follows along similar lines to those of Theorem (7.2), the only difference is that the real Kronecker canonical form is used for the establishment of the bases in part (ii). By slight modification of the conditions of Theorem (7.2) we have the following Corollaries.

Corollary (7.2): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $V\subset\mathbb{R}^n$ be a subspace and let $\dim V=d$. Necessary and sufficient conditions for V to be a prime o-(F,G)-e.d.s. and thus a prime ∞-(G,F)-e.d.s. are:

(i) $N_r(G)\cap V\neq 0$ and $N_r(F)\cap V=0$.

(ii) There exists a basis $B_o=\{\underline{x}_i^o, i\in\underset{\sim}{d}\}$ such that

$$G\underline{x}_i^o = F\underline{x}_{i-1}^o, \ i\in\underset{\sim}{d}, \ \underline{x}_o^o=\underline{0} \tag{7.17}$$

Proof

From the proof of Theorem (7.2) we have that $\mathcal{N}_r(F)\cap\mathcal{V}=0$ excludes the existence of c.m.i. for (FV,GV) pair, as well as the existence of infinite e.d. By setting $\alpha=0$ into (7.15a) and using Proposition (4.2), the result follows. □

Corollary (7.3): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\in\mathbb{R}^n$ be a subspace and let $\dim\mathcal{V}=d$. Necessary and sufficient conditions for $\mathcal{V}$ to be a prime ∞-(F,G)-e.d.s. and thus a prime o-(G,F)-e.d.s. are:

(i) $\mathcal{N}_r(F)\cap\mathcal{V}\neq 0$ and $\mathcal{N}_r(G)\cap\mathcal{V}=0$.

(ii) There exists a basis $\mathcal{B}_\infty=\{\underline{x}_i^\infty, i\in\underset{\sim}{d}\}$ such that

$$F\underline{x}_i^\infty = G\underline{x}_{i-1}^\infty,\ i\in\underset{\sim}{d},\ \underline{x}_o^\infty=\underline{0} \tag{7.18}$$

□

The bases $\mathcal{B}_\alpha,(\mathcal{B}_{\hat{\alpha}}),\mathcal{B}_{\sigma,\omega}(\mathcal{B}_{\hat{\sigma},\hat{\omega}}),\mathcal{B}_o,\mathcal{B}_\infty$ defined in the above results characterise the various types of prime (F,G) elementary divisor subspaces and and shall be referred to as characteristic bases. The above results clearly depend on the notion of a characteristic basis and in the following our attention is focussed on "basis free" characterisation of the prime, or simple, (F,G)-e.d.s.; some useful definitions and properties of special type pencils are considered first.

Definition (7.4): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$. The pair (F,G) will be called right nonsingular, (left nonsingular) if $C_r(F,G)$, $(C_\ell(F,G))$ is neutral. If (F,G) is right (left) nonsingular and $C_\ell(F,G)$ $(C_r(F,G))$ is either neutral, or simple, then (F,G) will becalled extended right regular (extended left regular). □

Clearly, if (F,G) is right nonsingular, then sF-G has no c.m.i. and if (F,G) is extended right regular, then it has no c.m.i. and no nonzero r.m.i.

Remark (7.4): If $sF-G\in\mathbb{R}^{m\times n}[s]$ is an extended right regular pencil, then $m\geq n$ and there exists an $R\in\mathbb{R}^{m\times m}$, $|R|\neq 0$, such that

$$R(sF-G)=\left[\begin{array}{c}0\\ \hdashline s\bar{F}-\bar{G}\end{array}\right]$$

where $s\bar{F}-\bar{G}$ is a regular restriction of $sF-G$. □

For the class of right nonsingular pencils we have the following result.

Proposition (7.6): Let $sF-G\in\mathbb{R}^{m\times n}[s]$ be a right nonsingular pencil and let $s\bar{F}-\bar{G}$ be a regular restriction. Then, for $\forall\alpha,\hat{\alpha}\in\mathbb{C}$ we have that

$$\mathcal{J}^r_\alpha(F,G)=\mathcal{J}^r_\alpha(\bar{F},\bar{G}) \text{ and } \mathcal{J}^r_{\hat{\alpha}}(G,F)=\mathcal{J}^r_{\hat{\alpha}}(\bar{G},\bar{F})$$

Furthermore, if (F,G) is extended right regular and then also for $\forall k$

$$\mathcal{N}_r(P^k_\alpha(F,G))=\mathcal{N}_r(P^k_\alpha(\bar{F},\bar{G})),\ \mathcal{N}_r(P^k_{\hat{\alpha}}(G,F))=\mathcal{N}_r(P^k_{\hat{\alpha}}(\bar{G},\bar{F}))$$ □

The latter result readily follows from the definitions, eqns (7.12) and Remark (7.4). Proposition (7.6) implies that all definitions, properties of right null spaces of $P^k_\alpha(F,G)$, $P^k_{\hat{\alpha}}(G,F)$ and structure of basis matrices for these null spaces derived in Chapter (4) for regular pencils also carry over to the case of right nonsingular pencils. Thus, notions such as the index of annihilation τ_α, the k-th generalised null space $\mathcal{M}^k_\alpha$ and of the sequences $\mathcal{J}_\alpha(F,G)$, defined for a regular pair (F,G) at $s=\alpha$ may also be used in exactly the same way for the case of right nonsingular pencils.

For a given pair $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, any subspace $\mathcal{V}\subset\mathbb{R}^n$ is characterised by the invariants of $(F,G)/\mathcal{V}$. Theorem (7.1) implies that the sequences $\mathcal{J}^r_\alpha(FV,GV)$, $\mathcal{J}^r_{\hat{\alpha}}(GV,FV)$, $\forall\alpha,\hat{\alpha}\in\mathbb{C}$ as well as $C_r(FV,GV)$, $C_\ell(FV,GV)$, where V is a basis of $\mathcal{V}$, are independent from the particular choice of the basis V and depend only on $\mathcal{V}$; thus, these sequences will be denoted by $\mathcal{J}^r_\alpha(F,G;\mathcal{V})$, $\mathcal{J}^r_{\hat{\alpha}}(G,F;\mathcal{V})$, $C_r(F,G;\mathcal{V})$, $C_\ell(F,G;\mathcal{V})$ and the root range by $\Phi(F,G;\mathcal{V})$. Note that if $\mathcal{V}$ is Φ-(F,G)-e.d.s., or $\mathcal{I}_{\mathcal{V}}=\{\mathcal{D}^s_{\alpha_i},\alpha_i\in\Phi(F,G;\mathcal{V})$ and possibly $\mathcal{I}_r\}$, then

$(F,G)/V$ is extended right regular, or right nonsingular respectively and the notions of the k-th generalised null space of $(F,G)/V$ at $s=\alpha$ (Definition (4.1)) and thus of the maximal generalised null space of $(F,G)/V$ at $s=\alpha$ are well defined; in such cases, the dimension of the maximal generalised null space of $(F,G)/V$ is independent of the particular choice of V, it will be denoted by $d_\alpha(F,G;V)$, and shall be called the <u>α-(F,G)-order</u> of V. A subspace V for which $(F,G)/V$ is right nonsingular will be referred to as an <u>extended-(F,G)-right regular subspace</u> (e-(F,G)-r.r.s.).

<u>Proposition (7.7)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, and $V\subset\mathbb{R}^n$ be a subspace.

(a) The subspace V is e-(F,G)-r.r.s. if and only if the sequence $C_r(F,G;V)$ is neutral.

(b) The subspace V is Φ-(F,G)-e.d.s. if and only if $C_r(F,G;V)$ is neutral and $C_\ell(F,G;V)$ is either neutral, or simple.

(c) A sufficient condition for V to be e-(F,G)-r.r.s. is that either $N_r(F)\cap V=0$ and/or $N_r(G)\cap V=0$. □

The proof is rather obvious and it is omitted. For simple Φ-(F,G)-e.d.s. we have the following result.

<u>Theorem (7.3)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $V\subset\mathbb{R}^n$ be a subspace, $\Phi(F,G;V)\neq 0$ be the root range of V and let $\dim V=d$.

(a) V is a simple α-(F,G)-e.d.s., $\alpha\in\mathbb{C}-\{0\}$, if and only if $\alpha\in\Phi(F,G;V)$ and

  (i) $N_r(F)\cap V=0$ and $N_r(G)\cap V=0$.

  (ii) If $\alpha\in\mathbb{R}$, then $d_\alpha(F,G;V)=d$. If $\alpha\in\mathbb{C}-\mathbb{R}$, then d is even and $d_\alpha(F,G;V)=d/2$.

(b) V is a simple ∞-(F,G)-e.d.s., if and only if

  (i) $N_r(F)\cap V\neq 0$ and $N_r(G)\cap V=0$.

  (ii) $o\in\Phi(G,F;V)$ and $d_o(G,F;V)=d$.

(c) V is a simple o-(F,G)-e.d.s., if and only if

(i) $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $\mathcal{N}_r(G)\cap\mathcal{V}\neq 0$.

(ii) $o\in\Phi(F,G;\mathcal{V})$ and $d_o(F,G;\mathcal{V})=d$.

Proof

(a) By Proposition (7.7), and conditions (i) it follows that $\mathcal{V}$ is e-(F,G)-r.r.s. and by Theorem (7.2) $\mathcal{I}_{\mathcal{V}}$ has no zero and infinite e.d. The pencil $(F,G)/\mathcal{V}$ is extended right nonsingular and thus since $\alpha\epsilon\Phi(G,F;\mathcal{V})$, $J^r_\alpha(F,G;\mathcal{V})$ is not neutral. If $\mathcal{I}_\alpha(\mathcal{V})=\{(d_i,\sigma_i)i\epsilon\underline{\rho}\}$ is the index set of $(F,G)/\mathcal{V}$ at $s=\alpha$, then $d_\alpha(F,G;\mathcal{V})=\Sigma d_i\sigma_i=d$ implies that $(F,G)/\mathcal{V}$ has only finite e.d. at $s=\alpha$ and possibly zero r.m.i. and thus sufficiency is established. The necessity of part (a) is obvious. The proof of the other parts is similar. □

The above result provides a basis free characterisation of simple Φ-(F,G)-e.d.s.; such a characterisation clearly depends on the properties of the $J^r_\alpha(F,G;\mathcal{V})$ sequences, which in turn define $d_\alpha(F,G;\mathcal{V})$.

Corollary (7.4): If $\mathcal{V}$ is an e-(F,G)-r.r.s., $\alpha\epsilon\Phi(F,G;\mathcal{V})$, $\mathcal{I}_\alpha(F,G;\mathcal{V})=\{(d_i,\sigma_i),i\epsilon\underline{\rho},d_1<\ldots<d_\rho\}$ is the index set of $(F,G)/\mathcal{V}$ at $s=\alpha$, and $\dim\mathcal{V}=d$, $d_\alpha(F,G;\mathcal{V})=\sum_{i=1}^{\rho}\sigma_i d_i$.

(i) If $\alpha\in\mathbb{R}\cup\{\infty\}$, then $d_\alpha(F,G;\mathcal{V})\leq d$ and equality holds, if and only if $\mathcal{V}$ is a simple α-(F,G)-e.d.s.

(ii) If $\alpha\epsilon\mathbb{C}-\mathbb{R}$, then $d_\alpha(F,G;\mathcal{V})\leq d/2$ and equality holds, if and only if $\mathcal{V}$ is a simple (α,α^*)-(F,G)-e.d.s. □

For the case of e-(F,G)-r.r.s. $\mathcal{V}$, it is clear that Definition (4.4) and Theorem (4.5) also apply for $(F,G)/\mathcal{V}$. In fact, if V is any basis matrix of $\mathcal{V}$, and $\mathcal{I}_\alpha(F,G;\mathcal{V})=\{(d_i,\sigma_i),i\epsilon\underline{\rho}\}$ is the index set of (FV,GV) at $s=\alpha$ $(\alpha\epsilon\mathbb{C}\cup\{\infty\})$, then the notion of a complete prime set of chains $\Sigma_\alpha(FV,GV)=$ $=\{\mathcal{S}_\alpha(d_1,\underline{v}_1^{d_1}),\ldots,\mathcal{S}_\alpha(d_1,\underline{v}_{\sigma_1}^{d_1});\ldots;\mathcal{S}_\alpha(d_\rho,\underline{v}_1^{d_\rho}),\ldots,\mathcal{S}_\alpha(d_\rho,\underline{v}_{\sigma_\rho}^{d_\rho})$, where

$\mathcal{B}_\alpha(FV,GV)=\{\underline{v}_1^{d_1},\dots,\underline{v}_{\sigma_1}^{d_1};\dots;\underline{v}_1^{d_\rho},\dots,\underline{v}_{\sigma_\rho}^{d_\rho}\}$ is a normal basis of generators of (FV,GV) at $s=\alpha$, is well defined. Every d_i-th order prime chain $\mathcal{S}_\alpha(d_i,\underline{v}_k^{d_i})=\{\underline{v}_{k,1}^{d_i},\dots,\underline{v}_{k,d_i}^{d_i}:\underline{v}_{k,1}^{d_i}=\underline{v}_k^{d_i}\}$ satisfies conditions (7.15a) or (7.15b) if $\alpha\in\mathbb{C}-\{0\}$, conditions (7.17) if $\alpha=0$ and conditions (7.18) if $\alpha=\infty$. Note that every chain $\mathcal{S}_\alpha(d_i,\underline{v}_k^{d_i})$ of (FV,GV) defines a chain of vectors in $\mathcal{V}$, if we set $\underline{x}_{k,j}^{d_i}=V\underline{v}_{k,j}^{d_i}$, $j\in\underset{\sim}{d}_i$; these chains of vectors in $\mathcal{V}$ will be denoted by $\mathcal{S}_\alpha(\underline{x}_k^{d_i};\mathcal{V})$, the set of such chains by $\Sigma_\alpha(F,G;\mathcal{V})$ and the corresponding set of generators by $\mathcal{B}_\alpha(F,G;\mathcal{V})$. The sets $\mathcal{B}_\alpha(F,G;\mathcal{V})$, $\mathcal{S}_\alpha(\underline{x}_k^{d_i};\mathcal{V})$ and $\Sigma_\alpha(F,G;\mathcal{V})$ are not uniquely defined, even for a given choice of basis matrix V and shall be referred to as an α-normal basis of generators in $\mathcal{V}$, an (α,d_i)-prime chain in $\mathcal{V}$ and an α-complete prime set of chains in $\mathcal{V}$ of (F,G) respectively. For a given choice of V basis matrix, the "Nested basis matrix approach", described in Chapter 4, may be applied on (FV,GV) for the computation of $\mathcal{B}_\alpha(F,G;\mathcal{V})$ and $\Sigma_\alpha(F,G;\mathcal{V})$ sets. For every chain $\mathcal{S}_\alpha(\underline{x}_k^{d_i};\mathcal{V})=\{\underline{x}_{k,j}^{d_i},j\in\underset{\sim}{d}_i,\underline{x}_{k,1}^{d_i}=\underline{x}_k^{d_i}\}$ the subspace defined by $\mathcal{V}(\underline{x}_k^{d_i})=\mathrm{sp}\,\mathcal{S}_\alpha(\underline{x}_k^{d_i};\mathcal{V})$ has dimension d_i, it is a subspace of $\mathcal{V}$ and shall be called an (α,d_i)-prime subspace in $\mathcal{V}$; the set of all such subspaces defined on $\Sigma_\alpha(F,G;\mathcal{V})$ will be denoted by $\mathcal{L}_\alpha(F,G;\mathcal{V})$ and shall be referred to as an α-complete prime set of subspaces in $\mathcal{V}$ of (F,G).

For an e-(F,G)-r.r.s. $\mathcal{V}$, apart from the sets $\mathcal{L}_\alpha(F,G;\mathcal{V})$ we may also define the notion of maximal generalised nullspace at $s=\alpha$. Thus, if V is a basis matrix, the notion of the maximal generalised nullspace of (FV,GV) at $s=\alpha$ is well defined, in a similar way to that of the regular case. Such a subspace, clearly depends on the choice of V and shall be denoted by $\mathcal{M}_\alpha^*(V)$; if X_V is a basis matrix of $\mathcal{M}_\alpha^*(V)$, then $\mathcal{N}_\alpha^*(V)=\mathrm{sp}\{VX_V\}$ is a subspace of $\mathcal{V}$ and shall be called an (α,V)-maximal generalised nullspace in $\mathcal{V}$ of (F,G) at $s=\alpha$ and its properties are discussed below.

Proposition (7.8): Let $\mathcal{V}$ be an e-(F,G)-r.r.s., $\dim\mathcal{V}=d$, $\Phi(F,G;\mathcal{V})$ be the root range, $\mathcal{I}_\alpha(F,G;\mathcal{V})=\{(d_i,\sigma_i),i\in\underset{\sim}{\rho}\}$ be the index set of $\mathcal{V}$ at $s=\alpha$ and let

V,V' be two basis matrices of $\mathcal{V}$. The following properties hold true:

(i) For any $\alpha\in\Phi(F,G;\mathcal{V})$, then any set $\Sigma_\alpha(F,G;\mathcal{V})$ and thus any set of subspaces $\mathcal{L}_\alpha(F,G;\mathcal{V})$ is linearly independent.

(ii) Let $\alpha\in\Phi(F,G;\mathcal{V})$, $\mathcal{L}_\alpha(F,G;\mathcal{V})=\{\mathcal{V}_k^{d_i}, i\in\underset{\sim}{\rho}, k\in\underset{\sim}{\sigma_i}\}$ be an α-complete prime set of subspaces derived for a given V and let $\mathcal{N}_\alpha^*(V)$ be the corresponding (α,V)-maximal generalised nullspace in $\mathcal{V}$. Then,

$$\mathcal{N}_\alpha^*(V) = \mathcal{V}_1^{d_1}\oplus\dots\oplus\mathcal{V}_{\sigma_1}^{d_1}\oplus\dots\oplus\mathcal{V}_1^{d_\rho}\oplus\dots\oplus\mathcal{V}_{\sigma_\rho}^{d_\rho} \tag{7.19}$$

and $\dim\mathcal{N}_\alpha^*(V)=d_\alpha(F,G;\mathcal{V})=\sum_{i=1}^{\rho}\sigma_i d_i$.

(iii) Let $\mathcal{N}_\alpha^*(V),\mathcal{N}_\alpha^*(V')$ be the maximal generalised nullspaces that correspond to two different basis matrices V,V'. Then $\mathcal{N}_\alpha^*(V)=\mathcal{N}_\alpha^*(V')$.

Proof

Part (i) and Part (ii) are straightforward consequences of Theorem (4.5). Let $N_\alpha^k, N_\alpha'^k$ be two nested basis matrices of $\mathcal{N}_r\{P_\alpha^k(FV,GV)\}$, $\mathcal{N}_r\{P_\alpha^k(FV',GV')\}$ respectively and let $V=V'Q, Q\in\mathbb{C}^{d\times d}, |Q|\neq 0$; it may be readily verified that $N_\alpha'^k=\text{diag}\{\underbrace{Q,\dots,Q}_{k}\}N_\alpha^k$. If τ is the index of annihilation of (FV,GV) and thus of (FV',GV') at s=α then

$$\tilde{X}'_\tau = [X_1'^1,\dots,X_\tau'^\tau] = Q[X_1^1,\dots,X_\tau^\tau] = Q\tilde{X}_\tau$$

where $\tilde{X}_\tau,\tilde{X}'_\tau$ are the last blocks of the naturally partitioned matrices $N_\alpha^\tau, N_\alpha'^\tau$. Then

$$\mathcal{N}_\alpha^*(V) = \text{sp}\{V\tilde{X}_\tau\} = \text{sp}\{V'QQ^{-1}\tilde{X}'_\tau\} = \text{sp}\{V'\tilde{X}'_\tau\} = \mathcal{N}_\alpha^*(V')$$

□

This result demonstrates that $\mathcal{N}_\alpha^*(V)$ is uniquely defined, and it is independent of the particular choice of V used for its definition; this subspace will be denoted by $\mathcal{N}_\alpha^*(\mathcal{V})$ and from now on shall be referred to as the α-maximal-(F,G)-generalised nullspace in $\mathcal{V}$. The above definitions and result clearly apply to the case of Φ-(F,G)-e.d.s.; for the case of simple

Φ-(F,G)-e.d.s. we may combine the characterisations provided by Theorem (7.3) and Proposition (7.8) to derive alternative criteria for such subspaces. In the following, if $\mathcal{W}\subset\mathbb{C}^n$ is a d-dimensional subspace, then by Re span $\mathcal{W}$ we shall denote the subspace of $\mathbb{R}^n$ which is spanned by the real and imaginary parts of the vectors in $\mathcal{W}$; clearly dim $\mathbb{R}$e span $\mathcal{W}$ =2d.

Corollary (7.5): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let dim $\mathcal{V}$=d. Then,

(a) $\mathcal{V}$ is a simple α-(F,G)-e.d.s., $\alpha\in\mathbb{R}-\{0\}$, if and only if

(i) $N_r(F)\cap\mathcal{V}=0$ and $N_r(G)\cap\mathcal{V}=0$.

(ii) $N_\alpha^*(\mathcal{V})=\mathcal{V}$.

(b) $\mathcal{V}$ is a simple (α,α^*)-(F,G)-e.d.s., $\alpha\in\mathbb{C}-\mathbb{R}$, if and only if

(i) $N_r(F)\cap\mathcal{V}=0$ and $N_r(G)\cap\mathcal{V}=0$.

(ii) $\mathbb{R}$e span $N_\alpha^*(\mathcal{V})=\mathcal{V}$.

(c) $\mathcal{V}$ is a simple o-(F,G)-e.d.s., if and only if

(i) $N_r(F)\cap\mathcal{V}=0$ and $N_r(G)\cap\mathcal{V}\neq 0$.

(ii) $N_o^*(\mathcal{V})=\mathcal{V}$.

(d) $\mathcal{V}$ is a simple ∞-(F,G)-e.d.s., if and only if

(i) $N_r(F)\cap\mathcal{V}=0$ and $N_r(G)\cap\mathcal{V}=0$.

(ii) $\mathcal{V}_\infty^*(\mathcal{V})=\mathcal{V}$. □

The above result is a straightforward consequence of Theorem (7.3), Proposition (7.8) and the definitions. An alternative characterisation of simple Φ-(F,G)-e.d.s. may be obtained by combining Corollary (7.5) and the decomposition result established by eqn.(7.19) of Proposition (7.8).

Corollary (7.6): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let dim $\mathcal{V}$=d.

(a) $\mathcal{V}$ is a simple α-(F,G)-e.d.s., or (α,α^*)-(F,G)-e.d.s., $\alpha\in\mathbb{C}-\{0\}$, if and only if

(i) $N_r(F)\cap\mathcal{V}=0$ and $N_r(G)\cap\mathcal{V}=0$.

(ii) If $L_{\alpha}(F,G;V)=\{V_{\alpha,k}^{d_{\alpha,i}}, i\in\underset{\sim}{\rho}_{\alpha}, k\in\underset{\sim}{\sigma}_{\alpha,i}\}$ is an α-complete prime set of subspaces in V, then

$$V = \mathbb{R}\text{e span } V_{\alpha,k}^{d_{\alpha,i}} \bigoplus_{i\in\rho_{\alpha}, k\in\sigma_{\alpha,i}} \tag{7.20a}$$

(b) V is a simple o-(F,G)-e.d.s., if and only if

(i) $N_r(F)\cap V=0$ and $N_r(G)\cap V\neq 0$.

(ii) If $L_o(F,G;V)=\{V_{o,k}^{d_{o,i}}, i\in\underset{\sim}{\rho}_o, k\in\underset{\sim}{\sigma}_{o,i}\}$ is a o-complete prime set of subspaces in V, then

$$V = V_{o,k}^{d_{o,i}} \bigoplus_{i\in\underset{\sim}{\rho}_o, k\in\underset{\sim}{\sigma}_{o,i}} \tag{7.20b}$$

(c) V is a simple ∞-(F,G)-e.d.s., if and only if

(i) $N_r(F)\cap V\neq 0$ and $N_r(G)\cap V=0$.

(ii) If $L_o(F,G;)=\{V_{\infty,k}^{d_{\infty,i}}, i\in\underset{\sim}{\rho}_{\infty}, k\in\underset{\sim}{\sigma}_{\infty,i}\}$ is an ∞-complete prime set of subspaces in V, then

$$V = V_{\infty,k}^{d_{\infty,i}} \bigoplus_{i\in\underset{\sim}{\rho}_{\infty}, k\in\underset{\sim}{\sigma}_{\infty,i}} \tag{7.20c}$$

□

<u>Remark (7.5)</u>: Let V_{α}^{d} be an (α,d)-prime subspace in V of (F,G). If $\alpha\in\mathbb{R}\cup\{\infty\}$, then V_{α}^{d} is a prime α-(F,G)-e.d.s. and $\dim V_{\alpha}^{d}=d$. If $\alpha\in\mathbb{C}-\{0\}$, then Re span V_{α}^{d} is a prime (α,α^*)-(F,G)-e.d.s. and dim $\mathbb{R}$e span $V_{\alpha}^{d}=2d$. Furthermore we have:

(i) If $\alpha\in\mathbb{R}-\{0\}$, there exist (α,d)-, and $(\hat{\alpha},d)$-prime chains of vectors $S_{\alpha}(\underline{x}_1;V)=\{\underline{x}_i; i\in\underset{\sim}{d}\}$, $S_{\hat{\alpha}}(\hat{\underline{x}}_1;V)=\{\hat{\underline{x}}_i; i\in\underset{\sim}{d}\}$, $\hat{\alpha}=\alpha^{-1}$ and $\underline{x}_1-\hat{\underline{x}}_1$, which define bases for V_{α}^{d} and satisfy conditions (7.15a),(7.15b) respectively.

(ii) If $\alpha=\sigma+j\omega\in\mathbb{C}-\mathbb{R}$, there exist (α,d)- and $(\hat{\alpha},d)$-prime chains of vectors $S_{\alpha}(\underline{x}_1;V)=\{\underline{x}_i=\underline{a}_i+j\underline{b}_i, i\in\underset{\sim}{d}\}$, $S_{\hat{\alpha}}(\hat{\underline{x}}_1;V)=\{\hat{\underline{x}}_i=\hat{\underline{a}}_i+j\hat{\underline{b}}_i, i\in\underset{\sim}{d}\}$, $\hat{\alpha}=\hat{\sigma}+j\hat{\omega}=\alpha^{-1}$ and $\underline{x}_1=\hat{\underline{x}}_1$, such that $\{(\underline{a}_i,\underline{b}_i), i\in\underset{\sim}{d}\}$, $\{(\hat{\underline{a}}_i,\hat{\underline{b}}_i), i\in\underset{\sim}{d}\}$ define bases for $\mathbb{R}$e span V_{α}^{d} and satisfy conditions (7.16a),(7.16b) correspondingly.

(iii) If α=0, there exists a (o,d)-prime chain of vectors $\mathcal{S}_o(\underline{x}_1;\mathcal{V})=\{\underline{x}_i; i\in\underset{\sim}{d}\}$, which defines a basis for $\mathcal{V}_o^d$ and satisfies conditions (7.17).

(iv) If α=∞, there exists an (∞,d)-prime chain of vectors $\mathcal{S}_\infty(\hat{\underline{x}}_1;\mathcal{V})=\{\hat{\underline{x}}_i; i\in\underset{\sim}{d}\}$, which defines a basis for $\mathcal{V}_\infty^d$ and satisfies conditions (7.18). □

By Corollary (7.6) and the above Remark it is clear that the notion of the characteristic basis for a simple α-, or (α,α*)-(F,G)-e.d.s. is well defined; such a basis is defined as the direct sum of the characteristic bases associated with the (α,d)-prime subspaces of the given α-, or (α,α*)-(F,G)-e.d.s.

Next, the case of $\mathcal{I}_c$-(F,G)-c.m.i.s. and $\mathcal{I}_r$-(F,G)-r.m.i.s. is examined. We first note that the definitions and results stated for the sequence $C_r(F,G)$, the right index of (F,G), σ_μ, and the maximal right annihilating space of (F,G), which were stated in Chapter (5) also apply to the case of the restriction pencil $(F,G)/\mathcal{V}$. Thus let $C_r(F,G;\mathcal{V})$ be the sequence associated with $(F,G)/\mathcal{V}$, $\sigma_\mu(\mathcal{V})$ the right index of $(F,G)/\mathcal{V}$, V a basis matrix of $\mathcal{V}$ and let $\bar{\mathbb{R}}^*(V)$ be the maximal right annihilating space of (FV,GV). If $\bar{R}_V$ is a basis matrix of $\bar{\mathcal{R}}^*(V)$, then the subspace $\mathcal{R}_V^*=\text{sp}\{V\bar{R}_V\}$ will be called <u>a V-maximal right annihilating space of (F,G) in $\mathcal{V}$</u>. Clearly, $C_r(F,G;\mathcal{V})$ and $\sigma_\mu(\mathcal{V})$ are uniquely determined by $\mathcal{V}$; the uniqueness of $\mathcal{R}_V^*$ is defined by the following result.

<u>Proposition (7.9)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let V be a basis matrix of $\mathcal{V}$. If $\mathcal{R}_V^*$ is the V-maximal right annihilating space of (F,G) in $\mathcal{V}$, then $\mathcal{R}_V^*$ is uniquely defined and it is independent from the particular choice of the basis matrix V.

<u>Proof</u>

Let σ_μ be the right index of $(F,G)/\mathcal{V}$ and let N_{σ_μ} be a basis matrix of $\mathcal{N}_r\{T_{\sigma_\mu}(FV,GV)\}$. By Corollary (5.8), $\mathcal{W}(N_{\sigma_\mu})=\bar{\mathcal{R}}^*(V)$ (N_{σ_μ} is naturally partitioned) and $T_{\sigma_\mu}(FV,GV)N_{\sigma_\mu}=0$. If $V=V'Q$, where $Q\in\mathbb{R}^{d\times d}$, $|Q|\neq 0$, $d=\dim \mathcal{V}$, then the last condition implies that $T_{\sigma_\mu}(FV',GV')\mathrm{diag}\{\underbrace{Q,\dots,Q}_{\sigma_\mu}\}N_{\sigma_\mu}=0$. Clearly, then $N'_{\sigma_\mu}=\mathrm{diag}\{\underbrace{Q,\dots,Q}_{\sigma_\mu}\}N_{\sigma_\mu}$ is a basis matrix for $\mathcal{N}_r\{T_{\sigma_\mu}(FV',GV')\}$ and thus $\mathcal{W}(N'_{\sigma_\mu})=\bar{\mathcal{R}}^*(V')$. From the definition of $\mathcal{W}(N_k)$ subspace (see Chapter (5)), it is readily shown that if $\bar{R}_V,\bar{R}_{V'}$ are basis matrices for $\bar{\mathcal{R}}^*(V),\bar{\mathcal{R}}^*(V')$ respectively, then $\bar{R}_{V'}=Q\bar{R}_V$ and thus

$$\mathcal{R}^*_V = \mathrm{sp}\{V\bar{R}_V\} = \mathrm{sp}\{V'QQ^{-1}\bar{R}_{V'}\} = \mathrm{sp}\{V'\bar{R}_{V'}\} = \mathcal{R}^*_{V'}$$

□

This result demonstrates that $\mathcal{R}^*_V$ is uniquely defined by $\mathcal{V}$ and it is independent from the particular choice of basis matrix V used for its definition; this subspace will be denoted by $\mathcal{R}^*(\mathcal{V})$ and from now on shall be referred to as the <u>maximal right annihilating space of (F,G) in $\mathcal{V}$</u>. Corollary (5.8) describes the properties of $\mathcal{R}^*(\mathcal{V})$ and establishes a procedure for its computation in terms of the properties of the naturally partitioned basis matrices N_{σ_μ} of $\mathcal{N}_r\{T_{\sigma_\mu}(FV,GV)\}$. The characterisation of $\mathcal{I}_c$-(F,G)-c.m.i.s. is given by the following result.

<u>Theorem (7.4)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace, $\dim\mathcal{V}=d$, $\mathcal{K}_r(F,G;\mathcal{V})=\{(\sigma_i,\rho_i),i\in\underline{\mu}\}$ be the right set of singularity of $(F,G)/\mathcal{V}$ and let $\mathcal{R}^*(\mathcal{V})$ be the maximal right annihilating space of (F,G) in $\mathcal{V}$. $\mathcal{V}$ is an $\mathcal{I}_c$-(F,G)-c.m.i.s., if and only if one of the following equivalent conditions hold true:

(i) $c_r(F,G;\mathcal{V}) = \sum_{i=1}^{\mu}\sigma_i\rho_i = d$.

(ii) $\mathcal{R}^*(\mathcal{V}) = \mathcal{V}$.

(iii) There exists a basis matrix $V=[\underline{x}_1,\dots,\underline{x}_d]$ of $\mathcal{V}$ such that

$$F\underline{x}_1=\underline{0},\ F\underline{x}_2=G\underline{x}_1,\dots,F\underline{x}_d=G\underline{x}_{d-1},\ G\underline{x}_d=\underline{0} \tag{7.21}$$

(iv) The subspace $\mathcal{V}$ may be written as

$$\mathcal{V} = \mathcal{R}^1_{\sigma_1}\oplus\dots\oplus\mathcal{R}^{\rho_1}_{\sigma_1}\oplus\dots\oplus\mathcal{R}^1_{\sigma_\mu}\oplus\dots\oplus\mathcal{R}^{\rho_\mu}_{\sigma_\mu} \tag{7.22}$$

where dim $\mathcal{R}^{j}_{\sigma_i}=\sigma_i$ $\forall j\epsilon\rho_i$ and every $\mathcal{R}^{j}_{\sigma_i}$ is a prime (σ_i-1)-(F,G)-c.m.i.s. □

This result is readily established by Proposition (7.9) and the results given in Chapter (5) for right singular pencils; in particular the subspaces $\mathcal{R}^{j}_{\sigma_i}$ are defined by Corollary (5.8). The number $c_r(F,G;\mathcal{V})$ is defined for every subspace $\mathcal{V}$ and every pair (F,G), it is an invariant of $\mathcal{V}$ with respect to (F,G) and shall be called the <u>right-(F,G)-order of $\mathcal{V}$</u>; $c_r(F,G;\mathcal{V})$ is readily computed from the sequence $\mathcal{C}_r(F,G;\mathcal{V})$, or the set $\mathcal{K}_r(F,G;\mathcal{V})$. In a similar manner we may introduce the sequence $\mathcal{C}_\ell(F,G;\mathcal{V})$, the left set of singularity $\mathcal{K}_\ell(F,G;\mathcal{V})=\{(\sigma_i',\rho_i'),i\epsilon\underset{\sim}{\mu}'\}$ of $(F,G)/\mathcal{V}$ and the number $c_\ell(F,G;\mathcal{V})=$ $=\sum_{i=1}^{\mu'}(\sigma_i'-1)\rho_i'$; the number $c_\ell(F,G;\mathcal{V})$ will be called the <u>left-(F,G)-order of $\mathcal{V}$</u>. The number $c_\ell(F,G;\mathcal{V})$ may be computed from $\mathcal{K}_\ell(F,G;\mathcal{V})$ in a similar manner to that given in Chapter (5). We close this section by giving a result characterising $\mathcal{I}_r$-(F,G)-r.m.i.s.

<u>Theorem (7.5)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace, dim $\mathcal{V}=d$, $\mathcal{K}_\ell(F,G;\mathcal{V})=\{(\sigma_i',\rho_i'),i\epsilon\underset{\sim}{\mu}'\}$ be the left set of singularity of $(F,G)/\mathcal{V}$ and let $c_\ell(F,G;\mathcal{V})$ be the left-(F,G)-order of $\mathcal{V}$. $\mathcal{V}$ is an $\mathcal{I}_r$-(F,G)-r.m.i.s. if and only if one of the following equivalent conditions hold true:

(i) $c_\ell(F,G;\mathcal{V})=d$.

(ii) $\mathcal{N}_r(F)\cap\mathcal{V}=0$, $\mathcal{N}_r(G)\cap\mathcal{V}=0$ and there is no proper subspace $\mathcal{V}'$ of $\mathcal{V}$ for which either $G\mathcal{V}'\subseteq F\mathcal{V}'$ and/or $F\mathcal{V}'\subseteq G\mathcal{V}'$.

(iii) The set $\mathcal{I}_{\mathcal{V}}=\{(\sigma_i'-1,\rho_i),i\epsilon\underset{\sim}{\mu}'\}$.

<u>Proof</u>

Part (i) follows by inspection of the Kronecker form of $(F,G)/\mathcal{V}$ and from the definition of $c_\ell(F,G;\mathcal{V})$. The conditions $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $\mathcal{N}_r(G)\cap\mathcal{V}=0$ exclude the presence of c.m.i., o-e.d. and ∞-e.d. from $\mathcal{I}_{\mathcal{V}}$; thus, if those two conditions are satisfied, then $\mathcal{I}_{\mathcal{V}}$ contains possibly finite nonzero e.d. and r.m.i. If there is a finite nonzero e.d., then by Theorem (7.2) and

Corollary (7.1), there exists a proper subspace $\mathcal{V}'$ of $\mathcal{V}$ for which $G\mathcal{V}'\subseteq F\mathcal{V}'$ and thus contradicts the other assumption. The sufficiency of part (ii) is obvious. Part (iii) is readily established from the definitions. □

The characterisations provided in this section provide the means for the introduction of some more general notions of invariant subspaces of the domain of (F,G), which will be considered next.

7.3.3. Classification of the subspaces of the domain of (F,G)

The properties of the characteristic bases associated with Φ-(F,G)-e.d.s. and the existence of the special bases characterising $\mathcal{I}_c$-(F,G)-c.m.i.s. (Theorem (7.4) parts (iii) and (iv)) potivate the following definition for subspaces of the domain of (F,G).

Definition (7.5): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let $\dim\mathcal{V}=d$.

(i) $\mathcal{V}$ will be called a (G,F)-invariant subspace ((G,F)-i.s.) if

$$G\mathcal{V}\subseteq F\mathcal{V} \tag{7.23a}$$

or equivalently, for any basis matrix V of $\mathcal{V}$, there exists an $\bar{A}\in\mathbb{R}^{d\times d}$ such that

$$GV = FV\bar{A} \tag{7.23b}$$

(ii) $\mathcal{V}$ will be called an (F,G)-invariant subspace ((F,G)-i.s.) if

$$F\mathcal{V}\subseteq G\mathcal{V} \tag{7.24a}$$

or equivalently, for any basis matrix V of $\mathcal{V}$, there exists an $\underline{A}\in\mathbb{R}^{d\times d}$ such that

$$FV = GV\underline{A} \tag{7.24b}$$

(iii) $\mathcal{V}$ will be called a complete-(F,G)-invariant subspace (c-(F,G)-i.s.) if

$$F\mathcal{V} = G\mathcal{V} \tag{7.25a}$$

or equivalently for any basis matrix V of $\mathcal{V}$ there exists a pair of matrices $\bar{A},\underline{A} \in \mathbb{R}^{d\times d}$ such that

$$GV = FV\bar{A} \text{ and } FV = GV\underline{A} \tag{7.25b}$$

□

The matrices $\bar{A},\underline{A}$ will be referred to as the V-(G,F)-, V-(F,G)-restrictions of the (G,F)-i.s., (F,G)-i.s. respectively and shall be denoted by $(G,F;\mathcal{V})/V$, $(F,G;\mathcal{V})/V$ correspondingly. The set of eigenvalues of $\bar{A},\underline{A}$ will be denoted by $\sigma(G,F;V)$, $\sigma(F,G;V)$ respectively and shall be referred to as the V-(G,F)-, V-(F,G)-spectrum of the (G,F)-i.s., (F,G)-i.s. correspondingly.

Using the notions of (G,F)-, (F,G)-invariance and complete (F,G)-invariance introduced above we may give the following geometric characterisations of the subspaces defined algebraically in Section (7.3.1).

Theorem (7.6): Let $(F,G) \in \mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let $\dim\mathcal{V}=d$.

(a) $\mathcal{V}$ is an ∞-proper, Φ_∞,-(F,G)-e.d.s., if and only if

(i) $\mathcal{N}_r(F)\cap\mathcal{V}=0$.

(ii) $G\mathcal{V}\subseteq F\mathcal{V}$.

(b) $\mathcal{V}$ is an o-proper,Φ_o,-(F,G)-e.d.s., if and only if

(i) $\mathcal{N}_r(G)\cap\mathcal{V}=0$.

(ii) $F\mathcal{V}\subseteq G\mathcal{V}$.

(c) $\mathcal{V}$ is a proper, Φ_p-(F,G)-e.d.s., if and only if

(i) $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $\mathcal{N}_r(G)\cap\mathcal{V}=0$.

(ii) $F\mathcal{V}=G\mathcal{V}$.

Proof

(a) The necessity of the conditions (i) and (ii) follows by inspection of the Kronecker form of $(F,G)/\mathcal{V}$, Theorem (7.2) and Corollaries (7.1) and (7.2). The sufficiency may be argued as follows: condition (i) excludes the existence of ∞-e.d. and c.m.i. in $\mathcal{I}_{\mathcal{V}}$; thus, the pencil $(F,G)/\mathcal{V}$ is right nonsingular and $\mathcal{V}$ is e-(F,G)-r.r.s. By condition (ii) we have that $GV=FV\bar{A}$ and thus if $\bar{A}=Q\bar{J}Q^{-1}$ is the Jordan decomposition of $\bar{A}$, then $GVQ=FVQJ$ and the vectors of $V_c=VQ$ define a characteristic basis for $\mathcal{V}$. Because $(F,G)/\mathcal{V}$ is right nonsingular, Proposition (4.3) applies for every $\alpha\in\Phi(F,G;\mathcal{V})$ and $\mathcal{I}_{\mathcal{V}}$ contains finite e.d., the degrees of which are defined by the dimensions of the Jordan blocks in $\bar{J}$. Clearly, the sum of the degrees of all e.d. is equal to d, the dimension of $\mathcal{V}$ and by inspection of the Kronecker form it follows that $\mathcal{V}$ has no nonzero r.m.i. The proofs of parts (b) and (c) follow along similar lines. □

It is evident from the above result that Φ_{∞},-(F,G)-e.d.s., Φ_{o},-(F,G)-e.d.s. and Φ_{p},-(F,G)-e.d.s. are special cases of (G,F)-i.s., (F,G)-i.s. and complete-(F,G)-i.s. respectively; the important characteristic for such subspaces is the uniqueness of the spectrum which is discussed next.

Corollary (7.7): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and let V be a basis matrix of $\mathcal{V}$. Then,

(a) If $\mathcal{V}$ is a Φ_{∞},-(F,G)-e.d.s., then for any V, $\sigma(G,F;V)$ is uniquely defined.

(b) If $\mathcal{V}$ is a Φ_{o},-(F,G)-e.d.s., then for any V, $\sigma(F,G;V)$ is uniquely defined.

(c) If $\mathcal{V}$ is a Φ_{p},-(F,G)-e.d.s., then for any V of the spectra $\sigma(G,F;V)$ and $\sigma(F,G;V)$ are uniquely defined. Furthermore, if $\bar{A}=(G,F;\mathcal{V})/V$, $\underline{A}=(F,G;\mathcal{V})/V$, then $\underline{A}=\bar{A}^{-1}$ and $\sigma(G,F;V)$ is the inverse of $\sigma(G,F;V)$.

Proof

(a) If $\mathcal{V}$ is Φ_∞,-(F,G)-e.d.s. and V,V' are two basis matrices, then $GV=FV\bar{A}$, $GV'=FV'\bar{A}'$. If $V'=VQ, Q\in\mathbb{R}^{d\times d}, |Q|\neq 0, d=\dim\mathcal{V}$, then from the last three relationships we have that $GV=FV\bar{A}=FVQ\bar{A}'Q^{-1}$ and thus $FV(\bar{A}-Q\bar{A}'Q^{-1})=0$. Since $\mathcal{N}_r(F)\cap\mathcal{V}=0$ it follows that $\mathcal{N}_r(FV)=0$ and thus the last condition implies $\bar{A}=Q\bar{A}^{-1}Q^{-1}$; the similarity of $\bar{A},\bar{A}'$ proves the result. Parts (b) and (c) are proved along similar lines. □

Because of the above property the subspaces of the Φ_∞,-(F,G)-i.s., Φ_o,-(F,G)-i.s., Φ_p,-(F,G)-i.s. type will be referred to as fixed spectrum invariant subspaces. Some interesting classes of assignable spectrum and partly fixed spectrum invariant subspaces will be introduced in the following. We first give some alternative characterisations of simple o-(F,G)-e.d.s., and ∞-(F,G)-e.d.s.

Corollary (7.8): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and let $\mathcal{V}$ be a subspace of $\mathbb{R}^n$.

(a) $\mathcal{V}$ is a simple o-(F,G)-e.d.s., if and only if
- (i) $\mathcal{N}_r(F)\cap\mathcal{V}=0$ and $G\mathcal{V}\subseteq F\mathcal{V}$.
- (ii) There is no proper subspace $\mathcal{V}'\subset\mathcal{V}$ for which $F\mathcal{V}'=G\mathcal{V}'$.

(b) $\mathcal{V}$ is a simple ∞-(F,G)-e.d.s., if and only if
- (i) $\mathcal{N}_r(G)\cap\mathcal{V}=0$ and $F\mathcal{V}\subseteq G\mathcal{V}$.
- (ii) There is no proper subspace $\mathcal{V}'\subset\mathcal{V}$ for which $F\mathcal{V}'=G\mathcal{V}'$.

Proof

(a) By conditions (i), $\mathcal{V}$ is a Φ_∞,-(F,G)-i.s. and thus $\mathcal{I}_\mathcal{V}$ is characterised by finite e.d. and possibly zero r.m.i. If there exist nonzero finite e.d. in $\mathcal{I}_\mathcal{V}$, then by the direct sum decomposition of $\mathcal{V}$ implied by the Kronecker canonical form and Theorem (7.6), there exists a proper subspace $\mathcal{V}'\subset\mathcal{V}$ for which $F\mathcal{V}'=G\mathcal{V}'$ and this contradicts assumption (ii); this proves the sufficiency. The necessity is obvious. Part (b) of the proof follows along similar lines. □

The notions of (F,G)-, (G,F)-invariance are also involved in the characterisation of $\mathcal{I}_c$-(F,G)-c.m.i.s. as it is shown below.

<u>Proposition (7.10):</u> Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and let $\mathcal{V}\subset\mathbb{R}^n$ be a subspace with $\dim\mathcal{V}=d$. $\mathcal{V}$ is an $\mathcal{I}_c$-(F,G)-c.m.i.s., if and only if

(i) $G\mathcal{V}=F\mathcal{V}$.

(ii) There exists a basis matrix V_c of $\mathcal{V}$ such that $(G,F;\mathcal{V})/V_c$ is a lower d-nillpotent and $(F,G;\mathcal{V})/V_c$ an upper d-nillpotent.

<u>Proof</u>

By Theorem (7.4), part (iii), there exists a basis matrix $V_c=[\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_d]$ such that conditions (7.21) are satisfied; clearly, eqns.(7.21) imply

$$GV_c = FV_c\bar{A}_o \text{ and } FV_c = GV_c\underline{A}_o \tag{7.26a}$$

where

$$\bar{A}_o = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}, \quad \underline{A}_o = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix} \in\mathbb{R}^{d\times d} \tag{7.26b}$$

Conditions (7.26a) imply that $F\mathcal{V}=G\mathcal{V}$ and the matrices $\underline{A}_o,\bar{A}_o$ are clearly d-nillpotent. The sufficiency is established by a mere reversion of the arguments. □

Next we examine the characterisation of (F,G)-i.s., (G,F)-i.s. and proper-(F,G)-i.s. in terms of the invariants of the restriction pencil. We first give the following result.

<u>Proposition (7.11)</u>: Let $(F,G) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$ and let $\mathcal{V} \in \mathbb{R}^n$ be a subspace.

(a) If $\mathcal{V}$ is an $\mathcal{I}_r$-(F,G)-r.m.i.s. $\mathcal{I}_r \neq \mathcal{I}_r^o$, then there exist vectors $\underline{v}, \underline{v}' \in \mathcal{V}$ such that $G\underline{v} \notin F\mathcal{V}$ and $F\underline{v}' \notin G\mathcal{V}$.

(b) If $\mathcal{V}$ is a simple ∞-(F,G)-e.d.s., then $F\mathcal{V} \subset G\mathcal{V}$; that is there exist vectors $\underline{v} \in \mathcal{V}$ such that $G\underline{v} \notin F\mathcal{V}$.

(c) If $\mathcal{V}$ is a simple o-(F,G)-e.d.s., then $G\mathcal{V} \subset F\mathcal{V}$; that is there exist vectors $\underline{v} \in \mathcal{V}$ such that $F\underline{v} \notin G\mathcal{V}$.

<u>Proof</u>

(a) There always exists a basis matrix V_k of $\mathcal{V}$ and an $R \in \mathbb{R}^{m\times m}$, $|R| \neq 0$ such that $(RFV_k, RGV_k) = (\tilde{F}V_k, \tilde{G}V_k)$ is the Kronecker form. For the sake of simplicity let us assume that

$$(\tilde{F}V_k, \tilde{G}V_k) = \left(\begin{bmatrix} 0 & 0 & 0 \\ \hline 1 & 0 & \multicolumn{1}{|c}{0} \\ 0 & 1 & \multicolumn{1}{|c}{0} \\ 0 & 0 & \multicolumn{1}{|c}{0} \\ \hline 0 & 0 & \multicolumn{1}{|c}{1} \\ 0 & 0 & \multicolumn{1}{|c}{0} \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ \hline 0 & 0 & \multicolumn{1}{|c}{0} \\ 1 & 0 & \multicolumn{1}{|c}{0} \\ 0 & 1 & \multicolumn{1}{|c}{0} \\ \hline 0 & 0 & \multicolumn{1}{|c}{0} \\ 0 & 0 & \multicolumn{1}{|c}{1} \end{bmatrix} \right) \tag{7.27a}$$

Choose a vector $\underline{v} = V_k\underline{e}_1 \in \mathcal{V}$, where $\underline{e}_1$ is the first standard basis vector of $\mathbb{R}^3$. By inspection of (7.27a), it is clear that the vector $\underline{\tilde{F}V_k e_1} \notin sp\{\tilde{G}V_k\}$; given that the columns of $\tilde{G}V_k$ are linearly independent, the last condition implies that $[\tilde{F}V_k\underline{e}_1, \tilde{G}V_k]$ has full rank and thus also the matrix

$$R^{-1}[\tilde{F}V_k\underline{e}_1, \tilde{G}V] = [FV_k\underline{e}_1, GV_k] \tag{7.27b}$$

has full rank. Condition (7.27b) implies that $FV_k\underline{e}_1 \notin G\mathcal{V}$. Similarly, it can be proved that $\tilde{G}V_k\underline{e}_2 \notin sp\{\tilde{F}V_k\}$ and thus $GV_k\underline{e}_2 \notin F\mathcal{V}$. The proof of parts (b) and (c) follows along similar lines.

□

Theorem (7.7): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let $\mathcal{I}_{\mathcal{V}}$ be the set of strict equivalence invariants of $(F,G)/\mathcal{V}$.

(a) If $\mathcal{V}$ is (G,F)-i.s., then the possible sets of invariants in $\mathcal{I}_{\mathcal{V}}$ are: $\mathcal{I}_{\mathcal{V}}=\{\mathcal{D}^s_o;\mathcal{D}^s_{\alpha_i},\alpha_i\in\mathbb{C}-\{0\};\mathcal{I}_c;\mathcal{I}^o_r\}$.

(b) If $\mathcal{V}$ is (F,G)-i.s., then the possible sets of invariants in $\mathcal{I}_{\mathcal{V}}$ are: $\mathcal{I}_{\mathcal{V}}=\{\mathcal{D}^s_\infty;\mathcal{D}^s_{\alpha_i},\alpha_i\in\mathbb{C}-\{0\};\mathcal{I}_c;\mathcal{I}^o_r\}$.

(c) If $\mathcal{V}$ is a complete-(F,G)-i.s., then the possible sets of invariants in $\mathcal{I}_{\mathcal{V}}$ are: $\mathcal{I}_{\mathcal{V}}=\{\mathcal{D}^s_{\alpha_i},\alpha_i\in\mathbb{C}-\{0\};\mathcal{I}_c;\mathcal{I}^o_r\}$.

Proof

Let us assume that $(F,G)/\mathcal{V}$ has in $\mathcal{I}_{\mathcal{V}}$ every possible type of strict equivalence invariant. There always exists an $R\in\mathbb{R}^{m\times m}$, $|R|\neq 0$ and a special basis matrix V_k of $\mathcal{V}$ such that $(RFV_k,RGV_k)=(\tilde{F}V_k,\tilde{G}V_k)$ is in Kronecker form. The Kronecker form implies a partitioning for V_k as $V_k=[V_\zeta,V_\varepsilon,V_\infty,V_o,V_\alpha]$, where $V_\zeta,V_\varepsilon,V_\infty,V_o,V_\alpha$ are the submatrices corresponding to the Kronecker blocks associated with r.m.i., c.m.i., ∞-e.d., o-e.d. and finite nonzero e.d. respectively; this partitioning also implies a direct sum decomposion for $\mathcal{V}$ as

$$\mathcal{V} = \mathcal{V}_\zeta\oplus\mathcal{V}_\varepsilon\oplus\mathcal{V}_o\oplus\mathcal{V}_\infty\oplus\mathcal{V}_\alpha \tag{7.28a}$$

where $\mathcal{V}_i=\mathrm{sp}\{V_i\}$ and i is any index from the set $\{\zeta,\varepsilon,o,\infty,\alpha\}=\mathcal{A}$. An obvious implication of the Kronecker canonical decomposition of $(\tilde{F}V_k,\tilde{G}V_k)$ is that if for some $i\in\mathcal{A}$ and some $\underline{a},\underline{a}'\in\mathbb{R}^d$, $d=\dim\mathcal{V}$, $\tilde{F}V_i\underline{a}\notin\tilde{G}\mathcal{V}_i$ and $\tilde{G}V_i\underline{a}'\notin\tilde{F}\mathcal{V}_i$, then $\tilde{F}V_i\underline{a}'\notin\tilde{G}\mathcal{V}$ and $\tilde{F}V_i\underline{a}'\notin\tilde{F}\mathcal{V}$, correspondingly. If we now denote by $(\tilde{G}V_k)$ and $(\tilde{F}V_k)$ sets of linearly independent columns of $\tilde{G}V_k$ and $\tilde{F}V_k$ respectively (in this case the nonzero columns of $\tilde{G}V_k$ and $\tilde{F}V_k$), then we have the relationships

$$\widetilde{F}V_{\underline{i}}\underline{a} \notin \widetilde{G}\mathcal{V} \iff [\widetilde{F}V_{\underline{i}}\underline{a},(\widetilde{G}V_k)] \text{ full rank} \qquad (7.28b)$$

$$\widetilde{G}V_{\underline{i}}\underline{a}' \notin \widetilde{F}\mathcal{V} \iff [\widetilde{G}V_{\underline{i}}\underline{a}',(\widetilde{F}V_k)] \text{ full rank} \qquad (7.28c)$$

By multiplying the right hand side matrices on the left by R^{-1} we have that

$$[FV_{\underline{i}}\underline{a},(GV_k)]\ ,\ [GV_{\underline{i}}\underline{a}',(FV_k)] \text{ full rank} \qquad (7.28c)$$

This analysis leads to the following conclusions:

(i) If for some $i\in A$ and some $\underline{v}\in\mathcal{V}_i$, $F\underline{v}\notin G\mathcal{V}_i$, then also $F\underline{v}\notin G\mathcal{V}$.

(ii) If for some $i\in A$ and some $\underline{v}'\in\mathcal{V}_i$, $G\underline{v}\notin F\mathcal{V}_i$, then also $G\underline{v}\notin F\mathcal{V}$.

Using the above two conclusions we may proceed with the proof of the result. Thus,

(a) Assume that $\mathcal{V}$ is (G,F)-i.s., all types of strict equivalence invariants are present and that $\mathcal{V}$ is decomposed as in (7.28a). If there exist nonzero r.m.i. in $\mathcal{I}_{\mathcal{V}}$, then by Proposition (7.11) there exists a vector $\underline{v}\in\mathcal{V}_\zeta$ such that $G\underline{v}\notin F\mathcal{V}_\zeta$; by the second of the above conclusions it follows that $G\underline{v}\notin F\mathcal{V}_\zeta$ also implies $G\underline{v}\notin F\mathcal{V}$ and this clearly violates the (G,F)-invariance property of $\mathcal{V}$. Thus, $\mathcal{I}_{\mathcal{V}}$ has nonzero r.m.i. Similarly, using Proposition (7.11) (part (b)) and the second conclusion, it may be proved that $\mathcal{I}_{\mathcal{V}}$ has no infinite e.d. The proofs of parts (b) and (c) follows along similar lines. $\square$

From the proof of Theorem (7.7) we have a decomposition result for a general subspace $\mathcal{V}\subset\mathbb{R}^n$ with respect to the pair (F,G). For a given subspace $\mathcal{V}$ we shall denote by $\mathcal{I}_c(\mathcal{V}),\mathcal{I}_r(\mathcal{V}),\mathcal{D}_o(\mathcal{V}),\mathcal{D}_\infty(\mathcal{V}),\mathcal{D}_\alpha(\mathcal{V}),\alpha\in\mathbb{C}-\{0\}$ the sets of c.m.i., r.m.i., zero-e.d., ∞-e.d., all finite nonzero e.d. of $(F,G)/\mathcal{V}$ respectively. The general decomposition result may be stated as follows [Kar. -1].

Proposition (7.12): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and let $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_{c}(\mathcal{V});\mathcal{I}_{r}(\mathcal{V});\mathcal{D}_{\infty}(\mathcal{V});\mathcal{D}_{o}(\mathcal{V});\mathcal{D}_{\alpha}(\mathcal{V})\}$. The subspace $\mathcal{V}$ may be expressed as

$$\mathcal{V} = \mathcal{V}_{\varepsilon}\oplus\mathcal{V}_{\zeta}\oplus\mathcal{V}_{\infty}\oplus\mathcal{V}_{o}\oplus\mathcal{V}_{\alpha} \tag{7.29a}$$

where for every of the subspaces in the decomposition we have the following properties:

$$\begin{aligned}&\mathcal{I}_{\mathcal{V}_{\varepsilon}} = \{\mathcal{I}_{c}(\mathcal{V});\mathcal{I}_{r}^{o}\},\ \mathcal{I}_{\mathcal{V}_{\zeta}} = \{\mathcal{I}_{r}(\mathcal{V})\}\\ &\mathcal{I}_{\mathcal{V}_{\infty}} = \{\mathcal{D}_{\infty}(\mathcal{V});\mathcal{I}_{r}^{o}\},\ \mathcal{I}_{\mathcal{V}_{o}} = \{\mathcal{D}_{o}(\mathcal{V});\mathcal{I}_{r}^{o}\},\ \mathcal{I}_{\mathcal{V}_{\alpha}} = \{\mathcal{D}_{\alpha}(\mathcal{V});\mathcal{I}_{r}^{o}\}\end{aligned} \tag{7.29b}$$

☐

From Theorem (7.7) and the decomposition result stated above we have the following decomposition of the (F,G)-i.s., (G,F)-i.s., and complete-(F,G)-i.s.

Corollary (7.9): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and let $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_{c}(\mathcal{V});\mathcal{I}_{r}(\mathcal{V});\mathcal{D}_{\infty}(\mathcal{V});\mathcal{D}_{o}(\mathcal{V});\mathcal{D}_{\alpha}(\mathcal{V})\}$ be the set of strict equivalence invariants of $(F,G)/\mathcal{V}$.

(a) If $G\mathcal{V}\subseteq F\mathcal{V}$, then $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_{c}(\mathcal{V});\mathcal{D}_{o}(\mathcal{V});\mathcal{D}_{\alpha}(\mathcal{V});\mathcal{I}_{r}^{o}\}$ and $\mathcal{V}$ may be decomposed as

$$\mathcal{V} = \mathcal{V}_{\varepsilon}\oplus\mathcal{V}_{o}\oplus\mathcal{V}_{\alpha} \tag{7.30}$$

where $\mathcal{V}_{\varepsilon},\mathcal{V}_{o},\mathcal{V}_{\alpha}$ are subspaces characterised by the (7.29b) properties.

(b) If $F\mathcal{V}\subseteq G\mathcal{V}$, then $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_{c}(\mathcal{V});\mathcal{D}_{\infty}(\mathcal{V});\mathcal{D}_{\alpha}(\mathcal{V});\mathcal{I}_{r}^{o}\}$ and $\mathcal{V}$ may be decomposed as

$$\mathcal{V} = \mathcal{V}_{\varepsilon}\oplus\mathcal{V}_{\infty}\oplus\mathcal{V}_{\alpha} \tag{7.31}$$

where $\mathcal{V}_{\varepsilon},\mathcal{V}_{\infty},\mathcal{V}_{\alpha}$ are subspaces characterised by the (7.29b) properties.

(c) If $G\mathcal{V}=F\mathcal{V}$, then $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_{c}(\mathcal{V});\mathcal{D}_{\alpha}(\mathcal{V});\mathcal{I}_{r}^{o}\}$ and $\mathcal{V}$ may be decomposed as

$$\mathcal{V} = \mathcal{V}_{\varepsilon}\oplus\mathcal{V}_{\alpha} \tag{7.32}$$

where $\mathcal{V}_{\varepsilon},\mathcal{V}_{\alpha}$ are subspaces characterised by the (7.29b) properties. ☐

The decomposition result given above provides the means for the characterisation of the invariant subspaces in terms of the associated spectra. Note that in the following the spectrum will be considered as the set of roots of the polynomials $\det\{\lambda I-(G,F;\mathcal{V})/V\}$, if $\mathcal{V}$ is (G,F)-invariant and $\det\{\lambda I-(F,G;\mathcal{V})/V\}$ if $\mathcal{V}$ is (F,G)-invariant.

Proposition (7.13): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and let $\dim\mathcal{V}=d$. $\mathcal{V}$ is an $\mathcal{I}_c$-(F,G)-c.m.i.s. if and only if the following conditions hold true:

(i) $G\mathcal{V}=F\mathcal{V}$.

(ii) If Λ is a symmetric set of d-complex numbers, then there exists a basis matrix V_Λ of $\mathcal{V}$ such that $\sigma(G,F;V_\Lambda)=\Lambda$.

Proof

If $\mathcal{V}$ is $\mathcal{I}_c$-(F,G)-c.m.i.s., then by Proposition (7.10) $G\mathcal{V}=F\mathcal{V}$. Furthermore, there exists a polynomial vector $\underline{x}(s)=[\underline{x}_d,\underline{x}_{d-1},\dots,\underline{x}_2,\underline{x}_1]\underline{e}_d(s)=X_d\underline{e}_d(s)$, where X_d is a basis matrix of $\mathcal{V}$ such that

$$\{sF-G\}X_d\underline{e}_d(s)=\underline{0}\quad\text{for }\forall s\in\mathbb{C}\tag{7.33a}$$

Assume first that $\Lambda=\{\lambda_i, i\in\underset{\sim}{d}\}$ is distinct symmetric set. Then $E_d(\Lambda)=[\underline{e}_d(\lambda_1),\dots,\underline{e}_d(\lambda_d)]$ has full rank (Vomdermode matrix) and thus $X_dE_d(\Lambda)$ has also full rank; furthermore, we have that

$$(\lambda_iF-G)X_d\underline{e}_d(\lambda_i)=\underline{0},\quad\forall\lambda_i\in\Lambda\tag{7.33b}$$

and this proves the necessity in the case of distinct spectrum.

Note that in the case where Λ is not distinct, we may define a modified matrix $E'_d(\Lambda)$, where now for repeated frequencies, we may define vectors $\underline{e}^k_d(\lambda_i)=\frac{d^k}{ds^k}\underline{e}_d(s)\Big|_{s=\lambda_i}$, $k=0,1,\dots,\nu$, where ν is the multiplicity of λ_i. Once more the corresponding matrix $E'_d(\Lambda)$ has full rank and the result readily follows.

The sufficiency may be argued as follows. By condition $F\mathcal{V}=G\mathcal{V}$ and Corollary (7.9), $\mathcal{V}=\mathcal{V}_\varepsilon\oplus\mathcal{V}_\alpha$; thus if we choose a basis matrix for V as $V=[V_\varepsilon,V_\alpha]$ where V_ε,V_α are basis matrices for $\mathcal{V}_\varepsilon,\mathcal{V}_\alpha$ respectively, then

$$GV_\varepsilon = FV_\varepsilon\bar{A} \tag{7.34a}$$

and because both $\mathcal{V}_\varepsilon$ and $\mathcal{V}_\alpha$ are (G,F)-invariant we have that

$$\bar{A} = \text{diag}\{\bar{A}_\varepsilon,\bar{A}_\alpha\} \tag{7.34b}$$

and thus $\sigma(G,F;V)=\{\sigma(\bar{A}_\varepsilon),\sigma(\bar{A}_\alpha)\}$. By Corollary (7.7), $\sigma(\bar{A}_\alpha)$ is fixed and thus if $\mathcal{V}_\alpha\neq 0$ the d-Λ spectrum cannot be assigned. Clearly, this contradicts assumption (ii) and the result is established. □

Corollary (7.10): Let $(F,G)\in\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace, $\dim\mathcal{V}=d$, $\mathcal{R}^*(\mathcal{V})$ be the maximal right annihilating space of (F,G) in $\mathcal{V}$ and let $c_r(F,G;\mathcal{V})=\mu$ be the corresponding order.

(a) $\mathcal{V}$ is a (G,F)-i.s., if and only if there exists a direct sum decomposition

$$\mathcal{V} = \mathcal{R}^*(\mathcal{V})\oplus\mathcal{V}_f \tag{7.34}$$

where $\mathcal{V}_f$ is (G,F)-i.s. and $\sigma(G,F;V_f)=\Lambda_f$ is a $(d-\mu)$-fixed spectrum for any basis matrix V_f of $\mathcal{V}_f$.

(b) $\mathcal{V}$ is an (F,G)-i.s., if and only if there exists a direct sum decomposition

$$\mathcal{V} = \mathcal{R}^*(\mathcal{V})\oplus\mathcal{V}_f'$$

where $\mathcal{V}_f'$ is an (F,G)-i.s., and $\sigma(F,G;V_f')=\Lambda_f'$ is a $(d-\mu)$-fixed spectrum for any basis matrix V_f' of $\mathcal{V}_f'$. □

The proof is rather obvious and it is omitted. The determination of maximal dimension (G,F)-, (F,G)-invariant subspaces in a given subspace

$\mathcal{V}$ of $\mathbb{R}^n$ is considered next. The case where $\mathcal{V}=\mathbb{R}^n$ is examined first and the results will be specialised next to the case of a general $\mathcal{V}$. We first define:

Definition (7.6): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$. We may define the following two sequences of subspaces of $\mathbb{R}^n$:

$$\mathcal{P}(G,F) \triangleq \{\mathcal{T}^o=\mathbb{R}^n,\ \mathcal{T}^{k+1}=G^{-1}(F\mathcal{T}^k),k\geq 0\} \tag{7.35}$$

$$\mathcal{P}(F,G) \triangleq \{\mathcal{J}^o=\mathbb{R}^n,\ \mathcal{J}^{k+1}=F^{-1}(G\mathcal{J}^k),k\geq 0\} \tag{7.36}$$

Note that in the (7.35),(7.36) sequences, G and F are considered as maps and thus if $\mathcal{W}\subseteq\mathbb{R}^m$ is a subspace, then

$$G^{-1}\mathcal{W} \triangleq \{\underline{x}\in\mathbb{R}^n:\ G\underline{x}\in\mathcal{W}\} \tag{7.37}$$

□

The sequences $\mathcal{P}(G,F)$,$\mathcal{P}(F,G)$ will be called (G,F)-, (F,G)-invariance generating sequences respectively and the terms used will be clarified by the study of their properties. Note that $\mathcal{P}(G,F)$ is a specialised form of a sequence of subspaces introduced recently by Bernhard [Ber. -1] and Aplevich [Apl. -1], whereas $\mathcal{P}(F,G)$ is the "dual" (swapping of the F,G maps) of the $\mathcal{P}(G,F)$; thus, the properties of $\mathcal{P}(F,G)$,$\mathcal{P}(G,F)$ readily follow from the results in [Apl. -1] (Theorems (11) and (12)). We may summarise those results in our present framework as follows:

Theorem (7.8): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and $\mathcal{P}(G,F)$ and $\mathcal{P}(F,G)$ be the sequences of subspaces of $\mathbb{R}^n$ defined by (7.35) and (7.36).

(a) The sequence $\mathcal{P}(G,F)$ is non-increasing and converges to an element $\mathcal{T}^*$ in at most n steps. $\mathcal{T}^*$ is the maximal (G,F)-invariant subspace in $\mathbb{R}^n$.

(b) The sequence $\mathcal{P}(F,G)$ is non-increasing and converges to an element $\mathcal{J}^*$ in at most n steps. $\mathcal{J}^*$ is the maximal (F,G)-invariant subspace in $\mathbb{R}^n$.

□

This result establishes the existence, the uniqueness and an iterative procedure for the computation of the subspaces $\mathcal{T}^*$ and $\mathcal{J}^*$ of $\mathbb{R}^n$. $\mathcal{T}^*,\mathcal{J}^*$ will be referred to as the maximal <u>(G,F)-,(F,G)-invariant subspaces</u> of the pair (F,G) respectively. The link of $\mathcal{T}^*,\mathcal{J}^*$ with the subspaces of a Kronecker decomposition of $\mathbb{R}^n$ defined by Proposition (7.12) is examined next. We first note that if $\mathcal{I}(F,G)=\{\mathcal{I}_r(F,G);\mathcal{I}_c(F,G);\mathcal{D}_o(F,G);\mathcal{D}_\alpha(F,G),\alpha\in\mathbb{C}-\{0\};\mathcal{D}_\infty(F,G)\}$ are the strict equivalence invariants of the pencil sF- G, then a decomposition of $\mathbb{R}^n$ defined by Proposition (7.12) by setting $\mathcal{V}=\mathbb{R}^n$

$$\mathbb{R}^n = \mathcal{V}_\zeta \oplus \mathcal{V}_\varepsilon \oplus \mathcal{V}_o \oplus \mathcal{V}_\alpha \oplus \mathcal{V}_\infty \tag{7.38a}$$

where for the subspaces involved we have the properties

$$\mathcal{I}_{\mathcal{V}_\zeta} = \{\mathcal{I}_r(F,G);\mathcal{I}_r^o\},\ \mathcal{I}_{\mathcal{V}_\varepsilon} = \{\mathcal{I}_c(F,G);\mathcal{I}_r^o\} \tag{7.38b}$$

$$\mathcal{I}_{\mathcal{V}_\infty} = \{\mathcal{D}_\infty(F,G);\mathcal{I}_r^o\},\ \mathcal{I}_{\mathcal{V}_o} = \{\mathcal{D}_o(F,G);\mathcal{I}_r^o\},\ \mathcal{I}_{\mathcal{V}_\alpha} = \{\mathcal{D}_\alpha(F,G);\mathcal{I}_r^o\}$$

will be referred to as a <u>Kronecker decomposition of $\mathcal{V}$ with respect to (F,G)</u> ((F,G)-k.d.). If a general subspace $\mathcal{V}$ of $\mathbb{R}^n$ is considered, then the above decomposition, defined by the Kronecker form of (F,G)/$\mathcal{V}$ will be also referred to as (F,G)-k.d. of $\mathcal{V}$.

<u>Corollary (7.11)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{I}(F,G)$ the set of invariants of sF−G and let

$$\mathbb{R}^n = \mathcal{V}_\zeta \oplus \mathcal{V}_\varepsilon \oplus \mathcal{V}_o \oplus \mathcal{V}_\alpha \oplus \mathcal{V}_\infty \tag{7.39a}$$

be a Kronecker decomposition of $\mathbb{R}^n$. If $\mathcal{T}^*$ is the maximal (G,F)-invariant subspace, $\mathcal{J}^*$ the maximal (F,G)-invariant subspace and $\mathcal{R}^*$ the maximal right annihilating space, then we have the following:

(i) $$\mathcal{R}^* = \mathcal{V}_\varepsilon \tag{7.39b}$$

(ii) $$\mathcal{T}^* = \mathcal{V}_\varepsilon \oplus \mathcal{V}_o \oplus \mathcal{V}_\alpha \tag{7.39c}$$

(iii) $$\mathcal{J}^* = \mathcal{V}_\varepsilon \oplus \mathcal{V}_\infty \oplus \mathcal{V}_\alpha \tag{7.39d}$$

(iv) The subspace defined by $\mathcal{W}^* = \mathcal{T}^* \cap \mathcal{J}^*$ is the maximal complete (F,G)-invariant subspace in $\mathbb{R}^n$ and may be expressed as

$$\mathcal{W}^* = \mathcal{T}^* \cap \mathcal{J}^* = \mathcal{V}_\varepsilon \oplus \mathcal{V}_\alpha \tag{7.39e}$$

Proof

Part (i) has already been established in Chapter (5). Let $\mathcal{T}' = \mathcal{V}_\varepsilon \oplus \mathcal{V}_o \oplus \mathcal{V}_\alpha$. and assyme that $\mathcal{T}' \subset \mathcal{T}^*$. We may choose a basis matrix T for $\mathcal{T}^*$ of the type $T = [T', \hat{T}]$ where T' is a basis matrix for $\mathcal{T}'$. By the (G,F)-invariance property of $\mathcal{T}'$ and $\mathcal{T}^*$ we have that

$$G[T', \hat{T}] = F[T', \hat{T}] \left[\begin{array}{c|c} \bar{A}_1 & \bar{A}_{12} \\ \hline 0 & \bar{A}_2 \end{array}\right] \tag{7.40}$$

Since this is nonzero, the eigenvalues of $\bar{A}_2$ are either fixed or assignable (appropriate choice of $\hat{T}$); it is readily shown however that both cases contradict the assumption about the invariants of sF-G, implicit in the (7.39a) decomposition. Thus, $\mathcal{T}' = \mathcal{T}^*$. The proof of the other parts is similar.

□

The subspace $\mathcal{W}^*$ defined above will be referred to as the maximal complete (F,G)-invariant subspace of the pair (F,G). Given that $\mathcal{R}^* = \mathcal{V}_\varepsilon$ and $\mathcal{W}^*$ are uniquely defined the subspace $\mathcal{V}_\alpha$ may be computed in the following way:

Computation of $\mathcal{V}_\alpha$: Construct a basis matrix W for $\mathcal{W}^*$ as: W=[R,P] where R is a basis matrix of $\mathcal{R}^* = \mathcal{V}_\varepsilon$. Then,

$$G[R,P] = F[R,P] \left[\begin{array}{c|c} \bar{A}_1 & \bar{A}_{12} \\ \hline 0 & \bar{A}_2 \end{array}\right] \tag{7.41a}$$

Then $\sigma(\bar{A}_2)=\sigma(G,F;\mathcal{V}_\alpha)$ and a basis matrix for $\mathcal{V}_\alpha$ may be constructed as follows: Let Q be real similarity transformation such that

$$\left[\begin{array}{c|c} \bar{A}_1 & \bar{A}_{12} \\ \hline 0 & \bar{A}_2 \end{array}\right] = Q \left[\begin{array}{c|c} \tilde{A}_1 & 0 \\ \hline 0 & \tilde{A}_2 \end{array}\right] Q^{-1} \;, \; Q = \left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline 0 & Q_{22} \end{array}\right] \tag{7.41b}$$

By (7.41a) and (7.41b) it follows that, if $W'=WQ=[RQ_{11},RQ_{12}+PQ_{21}]=[R',V]$, then

$$G[R',V] = F[R',V] \left[\begin{array}{c|c} \tilde{A}_1 & 0 \\ \hline 0 & \tilde{A}_2 \end{array}\right] \tag{7.41c}$$

and $V=RQ_{12}+PQ_{21}$ is a basis matrix for a $\mathcal{V}_\alpha$ subspace. □

The computation of the subspaces $\mathcal{V}_o,\mathcal{V}_\infty$ independently from the Kronecker form reduction is examined next. We first note that there are certain similarities between the subspaces $(\mathcal{V}_o,\mathcal{V}_\varepsilon)$ and $(\mathcal{V}_\infty,\mathcal{V}_\varepsilon)$; in fact by comparing conditions (7.17) and (7.21) first and then (7.18) and (7.21) it is clear that there exist common elements between the properties of the basis vectors of $(\mathcal{V}_o,\mathcal{V}_\varepsilon)$ pair and of the $(\mathcal{V}_\infty,\mathcal{V}_\varepsilon)$ pair. This observation the following definition of sequences of subspaces of $\mathbb{R}^n$.

<u>Definition (7.7)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$. We may define the following two sequences of subspaces of $\mathbb{R}^n$:

$$\mathcal{Q}(F,G) \triangleq \{\mathcal{K}^o=0,\mathcal{K}^{k+1}=F^{-1}(G\mathcal{K}^k),k\geq 0\} \tag{7.42}$$

$$\mathcal{Q}(G,F) \triangleq \{\mathcal{L}^o=0,\ \mathcal{L}^{k+1}=G^{-1}(F\mathcal{L}^k),k\geq 0\} \tag{7.43}$$

where again in (7.42) and (7.43) sequences, G and F are considered as maps and thus the symbols involved are interpreted as in (7.37). □

The sequence $\mathcal{Q}(F,G)$ has also been discussed in [Ber.-1] and [Apl.-1]; Q(G,F) is the dual sequence defined on the pair (G,F). The properties of $\mathcal{Q}(F,G)$ have been studied in [Apl.-1]; these properties are summarised below.

<u>Theorem (7.9)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and $\mathcal{Q}(F,G)$ and $\mathcal{Q}(G,F)$ be the sequences of subspaces of $\mathbb{R}^n$ defined by (7.42) and (7.43).

(a) The sequence $\mathcal{Q}(F,G)$ is nondecreasing and converges to a subspace $\mathcal{K}^*$ in at most n steps.

(b) The sequence $\mathcal{Q}(G,F)$ is nondecreasing and converges to a subspace $\mathcal{L}^*$ in at most n steps. □

<u>Corollary (7.12)</u>: Let $(\mathcal{K}^*)^\perp,(\mathcal{L}^*)^\perp$ be the orthogonal complements of the subspaces $\mathcal{K}^*,\mathcal{L}^*$ respectively of the pair (F,G)

(a) $\mathcal{K}^*$ has the following properties:

(i) $F\mathcal{K}^*\subset G\mathcal{K}^*$

(ii) $\mathcal{N}_r(F)\cap(\mathcal{K}^*)^\perp=0$, or in other words: F has full column rank on $\mathcal{K}^{*\perp}$, and $\mathcal{K}^*$ is the largest subspace satisfying (7.42).

(b) $\mathcal{L}^*$ has the following properties:

(i) $G\mathcal{L}^*\subset F\mathcal{L}^*$

(ii) $\mathcal{N}_r(G)\cap(\mathcal{L}^*)^\perp=0$, or in other words: G has full column rank on $\mathcal{L}^{*\perp}$, and $\mathcal{L}^*$ is the largest subspace satisfying (7.43). □

Parts (a) of Theorem (7.9) and Corollary (7.12) have been established in [Apl.-1], whereas parts (b) are the corresponding dual statements on the pair (G,F). The relationship of the subspaces $\mathcal{K}^*$ and $\mathcal{L}^*$ to the Kronecker decomposition of $\mathbb{R}^n$ is established by the following result.

<u>Corollary (7.13)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $I(F,G)$ the set of invariants of sF-G, $\mathcal{R}^*$ the maximal right annihilating space of (F,G) and let $\mathcal{K}^*,\mathcal{L}^*$ be the maximal subspaces of the sequences $\mathcal{Q}(F,G),\mathcal{Q}(G,F)$ respectively. If

$$\mathbb{R}^n=\mathcal{V}_\zeta\oplus\mathcal{V}_\varepsilon\oplus\mathcal{V}_o\oplus\mathcal{V}_\alpha\oplus\mathcal{V}_\infty \tag{7.44a}$$

is a Kronecker decomposition of $\mathbb{R}^n$, then we have the following properties:

(i) $\quad K^* = V_\varepsilon \oplus V_\infty$ (7.44b)

(ii) $\quad L^* = V_\varepsilon \oplus V_o$ (7.44c)

(iii) $\quad R^* = V_\varepsilon = K^* \cap L^*$ (7.44d)

□

Part (i) of the Corollary has been recently established by Loiseau [Loi.-1]; by using duality arguments on the pair (F,G), part (ii) follows. Part (ii) is an obvious consequence of (i) and (ii). The (F,G)-invariance of K^* and the (G,F)-invariance of L^* suggest that the procedure suggested for the computation of V_α may also be applied (after appropriate modification) for the computation of V_∞ and V_o. In fact, for the case of V_∞ construct a basis matrix K for K^* as K=[R,M], where R is a basis matrix for R^* and then use (F,G)-invariance and repeat the procedure as for V_α. The procedure for computing V_o is dual.

The subspace K^* will be referred to as the <u>maximal almost-(F,G)-right-annihilating space of (F,G)</u> (m.a.-(F,G)-r.a.s.) and L^* as the <u>maximal almost-(G,F)-right annihilating space</u> of (F,G) (m.a.-(G,F)-r.a.s.). These definitions are motivated by the similarities between the condition ((7.18),(7.21)) describing (V_∞,V_ε) and thus K^*, and the conditions ((7.17),(7.21)) describing (V_o,V_ε) and thus L^*. The importance of such subspaces will be further clarified in the following sections. The analysis so far reveals the following properties of the subspaces of the Kronecker decomposition of $\mathbb{R}^n$:

<u>Remark (7.6)</u>: The maximal subspaces of $\mathbb{R}^n$ which are defined on the pair (F,G), R^*,T^*,J^*,W^*,K^*,L^* are uniquely defined and they are related as follows:

(i) $\quad T^* = L^* \oplus V_\alpha$ and $J^* = K^* \oplus V_\alpha$ (7.45)

(ii) $\quad R^* = T^* \cap K^* = J^* \cap L^* = K^* \cap L^*$ (7.46)

Furthermore, the e.d. subspaces $\mathcal{V}_o,\mathcal{V}_\alpha,\mathcal{V}_\infty$ are not uniquely defined.

□

By combining Remark (7.6) and the spectrum characterisation of the subspaces $\mathcal{V}_\varepsilon,\mathcal{V}_o,\mathcal{V}_\infty$ we have:

Proposition (7.14): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and let $\mathcal{K}^*,\mathcal{L}^*$ be the maximal subspaces of the sequences $\mathcal{Q}(F,G),\mathcal{Q}(G,F)$ respectively.

(i) $\mathcal{K}^*$ is the maximal (F,G)-invariant subspace of $\mathbb{R}^n$ for which for some appropriate basis matrix K, $\sigma(F,G;K)=\{0,\dots,0\}$.

(ii) $\mathcal{L}^*$ is the maximal (G,F)-invariant subspace of $\mathbb{R}^n$ for which for some appropriate basis matrix K, $\sigma(G,F;L)=\{0,\dots,0\}$.

□

The results established for the sequences $\mathcal{P}(F,G),\mathcal{P}(G,F),\mathcal{Q}(F,G),\mathcal{Q}(G,F)$ of $\mathbb{R}^n$ and the corresponding maximal elements $\mathcal{J}^*,\mathcal{T}^*,\mathcal{K}^*,\mathcal{L}^*$ may be readily extended to the case of sequences of subspaces and corresponding maximal elements contained in a subspace $\mathcal{V}$ of $\mathbb{R}^n$. The essential tool for such an extension is clearly the $\mathcal{V}$-restricted ordered pair of maps $([f/\mathcal{V}]\ [g/\mathcal{V}])$, or equivalently the pair of matrices (FV,GV), where V is a basis matrix of $\mathcal{V}$. If $\dim\mathcal{V}=d$, then the sequences $\mathcal{P}(FV,GV),\mathcal{P}(GV,FV),\mathcal{Q}(FV,GV),\mathcal{Q}(GV,FV)$ are subspaces of $\mathbb{R}^d$; clearly then every such subspace $\bar{\mathcal{M}}$ with a basis matrix M defines a subspace $\mathcal{M}=\mathrm{sp}\{VM\}\subset\mathcal{V}$. In the following the above sequences will be denoted by $\mathcal{P}_{\mathcal{V}}(F,G),\mathcal{P}_{\mathcal{V}}(G,F),\mathcal{Q}_{\mathcal{V}}(F,G),\mathcal{Q}_{\mathcal{V}}(G,F)$ and the corresponding maximal elements by $\mathcal{J}^*(\mathcal{V}),\mathcal{T}^*(\mathcal{V}),\mathcal{K}^*(\mathcal{V}),\mathcal{L}^*(\mathcal{V})$. All results stated for $\mathbb{R}^n$, also hold true for the general $\mathcal{V}$.

We close this section by introducing some general families of invariant subspaces other than those of the (G,F)-, and/or (F,G)-invariant type. The following definitions also extend the key notion of spectrum. In the following, if ζ is a symmetric set of $\mathbb{C}$ (which might denote the spectrum of an invariant subspace), then $\hat{\zeta}$ denotes the inverse set; the inverse set ζ is defined in a similar manner to that of the inverse root range and its values in general are from $\mathbb{C}\cup\{\infty\}-\{0\}$.

Definition (7.8): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and let $\mathcal{V}\subset\mathbb{R}^n$ be a subspace. We may define the following:

(i) $\mathcal{V}$ will be called a $(\mathcal{W},\mathcal{U})$-partitioned invariant subspace (($\mathcal{W},\mathcal{U}$)-p.i.s.), if there exists a pair of subspaces $(\mathcal{W},\mathcal{U})$ such that

$$\mathcal{V}=\mathcal{W}\oplus\mathcal{U} \text{ where } G\mathcal{W}\subseteq F\mathcal{W} \text{ and } F\mathcal{U}\subseteq G\mathcal{U} \tag{7.47}$$

(ii) Let us assume that $\mathcal{V}$ is a $(\mathcal{W},\mathcal{U})$-p.i.s. If $G\mathcal{U}\not\subset F\mathcal{U}$, then $\mathcal{V}$ will be called an extended-(G,F)-invariant subspace (e-(G,F)-i.s.), and if $F\mathcal{W}\not\subset G\mathcal{W}$, then $\mathcal{V}$ will be called an extended-(F,G)-invariant subspace (e-(F,G)-i.s.). If $\mathcal{V}$ is both e-(G,F)-i.s. and e-(F,G)-i.s. (i.e. $G\mathcal{U}\not\subset F\mathcal{U}$ and $F\mathcal{W}\not\subset G\mathcal{W}$), then it will be called an extended complete-(F,G)-invariant subspace (e.c.-(F,G)-i.s.).

(iii) Let $\mathcal{V}$ be a $(\mathcal{W},\mathcal{U})$-p.i.s. and let $W,U,V=[W,U]$ be basis matrices for $\mathcal{W},\mathcal{U},\mathcal{V}$ respectively. Let us also denote by $\sigma(G,F;W),\sigma(F,G;U)$ and $\hat{\sigma}(G,F;W),\hat{\sigma}(F,G;U)$ the W-,U-spectra of $\mathcal{W},\mathcal{U}$ correspondingly. Then,

$$\sigma[(G,F);\mathcal{W}(W),\mathcal{U}(U)] \overset{\Delta}{=} \sigma(G,F;W)\cup\hat{\sigma}(F,G;U) \tag{7.48}$$

is defined as the $(\mathcal{W}(W),\mathcal{U}(U))$-(G,F)-spectrum of $\mathcal{V}$ and

$$\sigma[(F,G);\mathcal{W}(W),\mathcal{U}(U)] \overset{\Delta}{=} \hat{\sigma}(G,F;W)\cup\sigma(F,G;U) \tag{7.49}$$

is defined as the $(\mathcal{W}(U),\mathcal{U}(U))$-(F,G)-spectrum of $\mathcal{V}$. □

Note that in the above definitions, either $\mathcal{W}$ and/or $\mathcal{U}$ may be the zero subspace; in this case we define as the spectrum of 0, $\sigma(0)$, and the inverse spectrum, $\hat{\sigma}(0)$, to be the empty set $\emptyset$. In either of the (7.48), or (7.49) we shall adopt the following meaning for U: $\emptyset\cup\zeta\overset{\Delta}{=}\zeta$. The families of (G,F)-i.s., (F,G)-i.s. are clearly contained in the general family of $(\mathcal{W},\mathcal{U})$-p.i.s. and thus the definition of spectrum given above is a natural extension of the definitions given before. Using Theorem (7.7) and Corollary (7.9) we may classify the $(\mathcal{W},\mathcal{U})$-p.i.s. family by the following results.

Remark (7.7): If $\mathcal{V}$ is a $(\mathcal{W},\mathcal{U})$-p.i.s. with respect to the pair (F,G), then the possible sets of strict equivalence invariants of $(F,G)/\mathcal{V}$ are:

$$\mathcal{I}_{\mathcal{V}} = \{\mathcal{I}_c(\mathcal{V});\mathcal{I}_r^o;\mathcal{D}_\infty(\mathcal{V});\mathcal{D}_o(\mathcal{V});\mathcal{D}_\alpha(\mathcal{V})\} \tag{7.50a}$$

Furthermore, $\mathcal{V}$ may be decomposed as

$$\mathcal{V} = \mathcal{V}_c\oplus\mathcal{V}_\infty\oplus\mathcal{V}_o\oplus\mathcal{V}_\alpha \tag{7.50b}$$

where $\mathcal{V}_c,\mathcal{V}_\infty,\mathcal{V}_o,\mathcal{V}_\alpha$ are the subspaces defined by Proposition (7.12). □

Corollary (7.14): Let $(F,G) \in \mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let $\mathcal{I}_{\mathcal{V}}=\{\mathcal{I}_c(\mathcal{V});\mathcal{I}_r(\mathcal{V});\mathcal{D}_\infty(\mathcal{V});\mathcal{D}_o(\mathcal{V});\mathcal{D}_\alpha(\mathcal{V})\}$ denote the set of possible strict equivalence invariants of $(F,G)/\mathcal{V}$.

(a) $\mathcal{V}$ is an e-(G,F)-i.s., if and only if $\mathcal{I}_{\mathcal{V}}$ always contains a set $\mathcal{D}_\infty(\mathcal{V})$ and $\mathcal{V}$ has a decomposition

$$\mathcal{V} = \mathcal{W}\oplus\mathcal{V}_\infty \tag{7.51}$$

where $\mathcal{W}$ is (G,F)-i.s., $\mathcal{V}_\infty\neq 0$ and $\mathcal{I}_{\mathcal{V}_\infty}=\{\mathcal{D}_\infty(\mathcal{V});\mathcal{I}_r^o\}$.

(b) $\mathcal{V}$ is an e-(F,G)-i.s., if and only if $\mathcal{I}_{\mathcal{V}}$ always contains a set $\mathcal{D}_o(\mathcal{V})$ and $\mathcal{V}$ has a decomposition

$$\mathcal{V} = \mathcal{V}_o\oplus\mathcal{U} \tag{7.52}$$

where $\mathcal{U}$ is (F,G)-i.s., $\mathcal{V}_o\neq 0$ and $\mathcal{I}_{\mathcal{V}_o}=\{\mathcal{D}_o(\mathcal{V});\mathcal{I}_r^o\}$.

(v) $\mathcal{V}$ is an e.c.-(F,G)-i.s., if and only if $\mathcal{I}_{\mathcal{V}}$ always contains sets $\mathcal{D}_o(\mathcal{V}),\mathcal{D}_\infty(\mathcal{V})$ and $\mathcal{V}$ has a decomposition

$$\mathcal{V} = \mathcal{V}_o\oplus\mathcal{T}\oplus\mathcal{V}_\infty \tag{7.53}$$

where $\mathcal{T}$ is a c-(F,G)-i.s., $\mathcal{V}_o,\mathcal{V}_\infty \neq 0$ and $\mathcal{I}_{\mathcal{V}_o}=\{\mathcal{D}_o(\mathcal{V});\mathcal{I}_r^o\}$, $\mathcal{I}_{\mathcal{V}_\infty}=\{\mathcal{D}_\infty(\mathcal{V});\mathcal{I}_r^o\}$. □

Note that in the above result the subspaces $\mathcal{W},\mathcal{U},\mathcal{T}$ are not necessarily different from the zero subspace; thus, we may state the following remark.

<u>Remark (7.8):</u> For the families of (G,F)-i.s. and/or (F,G)-i.s. we have the following properties:

(a) If $\mathcal{V}$ is (G,F)-i.s., then $\mathcal{V}$ is an e-(F,G)-i.s., if and only if $\mathcal{I}_{\mathcal{V}}$ contains a set $\mathcal{D}_o(\mathcal{V})$.

(b) If $\mathcal{V}$ is (F,G)-i.s., then $\mathcal{V}$ is an e-(G,F)-i.s., if and only if $\mathcal{I}_{\mathcal{V}}$ contains a set $\mathcal{D}_\infty(\mathcal{V})$. □

The distinguishing feature of the families of e-(G,F)-i.s., e-(F,G)-i.s., and e.c.-(F,G)-i.s. is that the spectrum contains infinite frequencies; this is clarified by the following remark.

<u>Remark (7.9):</u> Let $\mathcal{V}$ be a $(\mathcal{W},\mathcal{U})$-p.i.s.

(a) If $\mathcal{V}$ is an e-(G,F)-i.s., then for any choice of a basis matrix V of $\mathcal{V}$, the $(\mathcal{W}(W),\mathcal{U}(U))$-(G,F)-spectrum of $\mathcal{V}$ always contains infinite frequencies.

(b) If $\mathcal{V}$ is an e-(F,G)-i.s., then for any choice of a basis matrix V of $\mathcal{V}$, the $(\mathcal{W}(W),\mathcal{U}(U))$-(F,G)-spectrum of $\mathcal{V}$ always contains infinite frequencies.

(c) If $\mathcal{V}$ is an e.c.-(F,G)-i.s., then for any choice of a basis matrix V of $\mathcal{V}$, the $(\mathcal{W}(W),\mathcal{U}(U))$-(G,F)-spectrum and the $(\mathcal{W}(W),\mathcal{U}(U))$-(F,G)-spectrum always contains infinite frequencies.

The number of infinite frequencies in the above spectra is defined by the dimensions of the subspaces $\mathcal{V}_\infty$, $\mathcal{V}_o$ and $(\mathcal{V}_o,\mathcal{V}_\infty)$ in the decompositions (7.51), (7.52) and (7.53) respectively. □

The presence of infinite frequencies in the spectra of e-(G,F)-i.s., e-(F,G)-i.s. and e.c.-(F,G)-i.s. may be further clarified by studying the asymptotic properties of appropriate (G,F)-, (F,G)-invariant subspaces. These results are based on the work of Jaffe and Karcanias [Jaf. & Kar.-1] on asymptotic transmission subspaces of linear systems; the key tool for the development of the asymptotic characterisation of e-(F,G)-i.s. and/or e-(G,F)-i.s. is the notion of a "canonical regular triple" which may be associated with any triple $(F,G;\mathcal{V})$, where $\mathcal{V} \subset \mathbb{R}^n$; this new notion will be developed in the following section and allows the reduction of many problems of analysis referred to the triple $(F,G;\mathcal{V})$ to standard problems of regular state space theory, which may be discussed on special type triples $(F',G';\mathcal{V}')$ [Jaf. & Kar.-1].

7.4 Canonical regular invariant realizations of $(F,G;\mathcal{V})$ triples

The algebraic, geometric and dynamic aspects of regular state space theory have been studied in terms of the matrix pencil theory [Kar.-1], [Jaf. & Kar.-1], [Kar. & MacB.-1] etc,; the particular characteristic of regular state space theory is that the pencil sF-G is entirely right singular, if the system is controllable [Kar.-2] and it is characterised by finite e.d. and c.m.i., if the system is uncontrollable [Kar. & MacB.-1]. The problem considered in this section is the investigation of the links between the properties of the general triple $(F,G;\mathcal{V})$, where $sF-G \in \mathbb{R}^{m\times n}[s]$ is a general pencil, $\mathcal{V} \subset \mathbb{R}^n$ is a subspace with $\dim \mathcal{V} = d$ and the standard theory of regular state space systems described by the triple of maps S(A,B,C) describing a linear system. The results discussed in this section provide the basis for the study of dynamic aspects of autonomous generalized differential systems presented in the final section of this chapter. We first define:

Definition (7.9): Let $(F,G;\mathcal{V})$ be a triple, where $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and $\dim\mathcal{V}=d$.

(i) A triple $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$, where $(\tilde{F},\tilde{G})\in\mathbb{R}^{m\times\tilde{n}}\times\mathbb{R}^{m\times\tilde{n}}$, $\tilde{\mathcal{V}}\subset\mathbb{R}^{\tilde{n}}$ is a subspace with $\dim\tilde{\mathcal{V}}=d$, will be called a regular invariant realization (r.i.r.) of the triple $(F,G;\mathcal{V})$, if $s\tilde{F}-\tilde{G}$ is characterised by c.m.i. and possibly finite e.d. and the restriction pencils $(F,G)/\mathcal{V}$, $(\tilde{F},\tilde{G})/\tilde{\mathcal{V}}$ are strict equivalent. If $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ is a r.i.r. and $s\tilde{F}-\tilde{G}$ is entirely right singular, then it will be called a proper regular invariant realization (p.r.i.r.); finally, if $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ is a r.i.r. and $\tilde{n}$ is minimal, then it will be called a minimal regular invariant realization (m.r.i.r.).

(ii) Let $S(A,B)=\{\underline{x}=A\underline{x}+B\underline{u},A\in\mathbb{R}^{\tilde{n}\times\tilde{n}},B\in\mathbb{R}^{\tilde{n}\times\tilde{\ell}},\tilde{\ell}<\tilde{n},\operatorname{rank}B=\tilde{\ell}\}$ be a linear system, $\mathcal{X}\equiv\mathbb{R}^{\tilde{n}}$ be the state space of $S(A,B)$, $\tilde{\mathcal{V}}\in\mathbb{R}^{\tilde{n}}$ be a subspace with $\dim\tilde{\mathcal{V}}=d$ and let $sN-NA\in\mathbb{R}^{(\tilde{n}-\tilde{\ell})\times\tilde{n}}[s]$ be the restricted input-state pencil of $S(A,B)$. The pair of the system $S(A,B)$ and the subspace $\tilde{\mathcal{V}}$, $(S(A,B),\tilde{\mathcal{V}})$, will be called a regular forced invariant realization (r.f.i.r.) of $(F,G;\mathcal{V})$ and shall be denoted by $S_{i.r.}(A,B,\tilde{\mathcal{V}})$, if $(N,NA;\tilde{\mathcal{V}})$ is a r.i.r. of $(F,G;\tilde{\mathcal{V}})$. If $(N,NA;\tilde{\mathcal{V}})$ is p.r.i.r., or m.r.i.r., then $S_{i.r.}(A,B,\tilde{\mathcal{V}})$ will be called respectively a proper regular forced invariant realization (p.r.f.i.r.), or minimal regular forced invariant realization (m.r.f.i.r.) respectively of $(F,G;\mathcal{V})$. □

The existence of regular invariant realizations $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ and regular forced invariant realizations $S_{i.r.}(A,B,\tilde{\mathcal{V}})$ is examined next; we first state the following result, which is used in our present study.

Lemma (7.2): Let $S(A,B,C)=\{\underline{x}=A\underline{x}+B\underline{u},\underline{y}=C\underline{x},A\in\mathbb{R}^{n\times n},B\in\mathbb{R}^{n\times\ell},C\in\mathbb{R}^{m\times n},\ell,m<n$, $\operatorname{rank}B=\ell$, $\operatorname{rank}C=m\}$ be a linear system, $P(s)$ be the Rosenbrock's system matrix pencil, $Z(s)=sNM-NAM$ be the zero pencil and $\mathcal{E}_{K}(A,B,C)$ be the

Kronecker orbit of $\mathcal{S}(A,B,C)$.

(i) If $\mathcal{E}(Z(s))$ is the zet of zero pencils which correspond to systems $\mathcal{S}(A',B',C')\in\mathcal{E}_{K}(A,B,C)$, then $\mathcal{E}(Z(s))$ is a strict equivalence class.

(ii) If $\Psi_P=\{(s-\lambda_i)^{\tau_i},i\in\underset{\sim}{\rho};\hat{s}^{q_i},i\in\underset{\sim}{\mu},q_i\geq 2;\varepsilon_i>0,i\in\underset{\sim}{p};\eta_i>0,i\in\underset{\sim}{t}\}$ is the set of strict equivalence invariants of P(s), then the corresponding set of strict equivalence invaraints of $\mathcal{E}(Z(s))$ is

$$\Psi_Z=\{(s-\lambda_i)^{\tau_i},i\in\underset{\sim}{\rho};\hat{s}^{q_i'},q_i'=q_i=2,i\in\underset{\sim}{\mu};\varepsilon_i'=\varepsilon_i-1,i\in\underset{\sim}{p};\\ \eta_i'=\eta_i-1,i\in\underset{\sim}{t}\} \tag{7.54}$$

where in Ψ_Z, those q_i' for which $q_i'=0$ are omitted. □

The above result has been established in [Kar. & MacB.-1] and describes the "plus two" property for the i.e.d. and the "plus one" property for the c.m.i. and r.m.i. of the pencils P(s) and Z(s) associated with $\mathcal{S}(A,B,C)$; a proof of this result is given in[Kar.&Hay.,2]. Before we proceed with the study of r.i.r. of an $(F,G;\mathcal{V})$ triple we note that if $\mathcal{V}=\mathbb{R}^n$, the domain of (F,G), then the r.i.r. of $(F,G;\mathbb{R}^n)$ will be simply referred to as r.i.r. of the pair (F,G). Using Lemma (7.2) we may state the following result.

<u>Theorem (7.10)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, sF-G the associated pencil, $\mathcal{I}(F,G)=\{(s-\lambda_i)^{\tau_i},i\in\underset{\sim}{\rho};\hat{s}^{q_i},i\in\underset{\sim}{\mu};\varepsilon_i,i\in\underset{\sim}{p};\eta_i,i\in\underset{\sim}{t}\}$ the set of strict equivalence invariants of sF-G and let $\bar{n}_f=\sum_{i=1}^{\rho}\tau_i$, $\bar{n}_\infty=\sum_{i=1}^{\mu}q_i$, $\bar{n}_\varepsilon=\sum_{i=1}^{p}\varepsilon_i$ and $\bar{n}_\eta=\sum_{i=1}^{t}\eta_i$.

(i) There exists a Kronecker orbit $\mathcal{E}_K(A,B,C)$ of $\mathcal{S}(A,B,C)$ systems with an associated strict equivalence class of zero pencils $\mathcal{E}(Z(s))$ for which $sF-G\in\mathcal{E}(Z(s))$.

(ii) If $\mathcal{S}(A,B,C)\in\mathcal{E}_K(A,B,C)$ and P(s) is the system matrix pencil of $\mathcal{S}(A,B,C)$, then the set of strict equivalence invariants of P(s) is given by

$$\Psi_P = \{(s-\lambda_i)^{\tau_i}, i\in\underset{\sim}{\rho}; \hat{s}^{\bar{q}_i}, i\in\underset{\sim}{\mu}, \bar{q}_i=q_i+2; \bar{\varepsilon}_i=\varepsilon_i+1, i\in\underset{\sim}{\rho};$$

$$\bar{\eta}_i=\eta_i+1, i\in\underset{\sim}{t}\} \tag{7.55}$$

(iii) If $n_f=\bar{n}_f$, $n_\infty=\bar{n}_\infty+\mu$, $n_\varepsilon=\bar{n}_\varepsilon+p$, $n_\eta=\bar{n}_\eta+t$ and $\mathcal{S}(A,B,C)\in\mathcal{E}_K(A,B,C)$ then $A\in\mathbb{R}^{\tilde{n}\times\tilde{n}}$, $B\in\mathbb{R}^{\tilde{n}\times\tilde{\ell}}$, $C\in\mathbb{R}^{\tilde{m}\times\tilde{n}}$, where $\tilde{n}=n_f+n_\infty+n_\varepsilon+n_\eta$, $\tilde{\ell}=p+\mu$, $\tilde{m}=t+\mu$, $\tilde{n}-\tilde{\ell}=m$ and $\tilde{n}-\tilde{m}=n$. Furthermore, $\tilde{n}$ is the minimal dimension of $\mathcal{S}(A,B,C)$ systems for which parts (i) and (ii) hold true. □

The proof of the above result follows immediately from Lemma (5.2) by a mere reversion of the steps of the proof given in Appendix (I). The existence and construction of r.i.r. of a pair (F,G) and of triples (F,G;$\mathcal{V}$) may now be examined using the results provided by Theorem (7.10). The orbit $\mathcal{E}_K(A,B,C)$ established by the previous result will be called the <u>natural Kronecker orbit</u> (n.k.o.) of the pair (F,G). The orbit $\mathcal{E}_K(A,B,C)$ of (F,G) is characterised by a canonical element, the Kronecker canonical form, [Th.-1],[Mor.-1],[Kar. & MacB.-1], and it is defined next.

<u>Corollary (7.15)</u>: The natural Kronecker orbit $\mathcal{E}_K(A,B,C)$ of the pair (F,G) is characterised by a canonical triple, (A_K,B_K,C_K), the Kronecker canonical form, which is defined by

$$A_K=\mathrm{diag}\{A_\varepsilon,A_\eta,A_\infty,A_f\},\ C_K=\begin{bmatrix}0 & C_\eta & 0 & 0\\ 0 & 0 & C_\infty & 0\end{bmatrix},\ B_K=\begin{bmatrix}B_\varepsilon & 0\\ 0 & 0\\ 0 & B_\infty\\ 0 & 0\end{bmatrix} \tag{7.56a}$$

$$A_\varepsilon=\mathrm{diag}\{H(\bar{\varepsilon}_i), i\in\underset{\sim}{\rho}\}\in\mathbb{R}^{n_\varepsilon\times n_\varepsilon},\ A_\eta=\mathrm{diag}\{H(\bar{\eta}_i), i\in\underset{\sim}{t}\}\in\mathbb{R}^{n_\eta\times n_\eta}$$

$$A_\infty=\mathrm{diag}\{H(\tilde{q}_i),\tilde{q}_i=q_i+1, i\in\underset{\sim}{\mu}\}\in\mathbb{R}^{n_\infty\times n_\infty},\ A_f=\mathrm{diag}\{J(\lambda_i,\tau_i), i\in\underset{\sim}{\rho}\}\in\mathbb{R}^{n_f\times n_f}$$

$$B_\varepsilon=\text{block-diag}\{\underline{w}(\bar{\varepsilon}_i), i\in\underset{\sim}{\rho}\}\in\mathbb{R}^{n_\varepsilon\times p},\ B_\infty=\text{block-diag}\{\underline{w}(\tilde{q}_i), i\in\underset{\sim}{\mu}\}\in\mathbb{R}^{n_\infty\times\mu} \tag{7.56b}$$

$$C_\eta=\text{block-diag}\{\underline{v}^t(\bar{\eta}_i), i\in\underset{\sim}{t}\}\in\mathbb{R}^{t\times n_\eta},\ C_\infty=\text{block-diag}\{\underline{v}^t(\tilde{q}_i), i\in\underset{\sim}{\mu}\}\in\mathbb{R}^{\mu\times n_\infty}$$

where $J(\lambda,\tau)$ is a $\tau\times\tau$ Jordan block that corresponds to $s=\lambda$, $H(f)=J(0,f)$ and $\underline{w}(k)=[0,\dots,0,1]^t$, $\underline{v}^t(k')=[1,0,\dots,0]$, where $\underline{w}(k)\in\mathbb{R}^{k\times 1}$ and $\underline{v}^t(k')\in\mathbb{R}^{1\times k'}$. □

Remark (7.10): Let (A_K,B_K,C_K) be the Kronecker form of the natural Kronecker orbit $\mathcal{E}_K(A,B,C)$ of (F,G). The general element (A,B,C) of $\mathcal{E}_K(A,B,C)$ may be expressed as

$$A = T\{A_K+B_KL+KC_K\}T^{-1},\ B = T^{-1}B_KR,\ C = GC_KT \tag{7.57}$$

where $T\in\mathbb{R}^{\tilde{n}\times\tilde{n}}$, $R\in\mathbb{R}^{\tilde{\ell}\times\tilde{\ell}}$, $G\in\mathbb{R}^{\tilde{m}\times\tilde{m}}$, $|T|,|R|,|G|\neq 0$, otherwise arbitrary, and $L\in\mathbb{R}^{\tilde{\ell}\times\tilde{n}}$, $K\in\mathbb{R}^{\tilde{n}\times\tilde{m}}$ arbitrary. □

The number $\tilde{n}=n_f+n_\infty+n_\varepsilon+n_\eta$ defined in Theorem (7.10) will be called the Kronecker order of the pair (F,G).

Corollary (7.16): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $sF-G$ the associated pencil and let $\mathcal{I}(F,G)$ have the general form of Theorem (7.10). For any $\mathcal{S}(A,B,C)\in$ $\in\mathcal{E}_K(A,B,C)$ we have that:

(i) If N is a left annihilator of B and $\tilde{\mathcal{V}}=\mathcal{N}_r(C)$, then the triple $(N,NA;\tilde{\mathcal{V}})$ is a m.r.i.r. of (F,G) and $\mathcal{S}_{i.r.}(A,B,\tilde{\mathcal{V}})$ is a m.r.f.i.r. of (F,G).

(ii) There exist elements $\mathcal{S}(A',B',C')\in\mathcal{E}_K(A,B,C)$ with (A',B') controllable pairs such that the corresponding triples $(N',N'A';\tilde{\mathcal{V}}')$ define proper minimal regular invariant relaizations (p.m.r.i.r.) of (F,G) and $\mathcal{S}_{i.r.}(A',B';\tilde{\mathcal{V}}')$ proper minimal regular forced invariant realizations (p.m.r.f.i.r.) of (F,G). □

The second part of the above result follows from the fact that the use of output injection (matrix K in (7.57)) affects the controllability properties of $(A+KC,B)$ pair ([Kar. & MacB.-1]). An interesting special case arises when the pencil $sF-G$ has no r.m.i. and no i.e.d.; then the orbit $\mathcal{E}_K(A,B,C)$ is characterised by the property that $C=0$ and thus $\tilde{\mathcal{V}}=\mathbb{R}^{\tilde{n}}$. A pencil characterised by c.m.i. and possibly f.e.d. will be referred to as a <u>prime pencil</u> and the associate pair will also be called prime. For such pencils we have the following simplified result.

<u>Corollary (7.17)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ be a prime pair and let $\mathcal{I}(F,G)=\{(s-\lambda_i)^{\tau_i}, i\in\underset{\sim}{\rho}; \varepsilon_i, i\in\underset{\sim}{p}\}$.

(i) There exists a Kronecker orbit $\mathcal{E}_K(A,B)$ of $\mathcal{S}(A,B)$ systems with an associated strict equivalence class of restricted input-state pencils $\mathcal{E}(S_i(s))$ for which $sF-G\in\mathcal{E}(S_i(s))$.

(ii) If $\mathcal{S}(A,B)\in\mathcal{E}_K(A,B)$ and $C(s)=[sI-A,-B]$ is the input-state pencil, then the set of strict equivalence invariants of $C(s)$ is given by

$$\Psi_C = \{(s-\lambda_i)^{\tau}, i\in\underset{\sim}{\rho}; \bar{\varepsilon}_i=\varepsilon_i+1, i\in\underset{\sim}{p}\} \tag{7.58}$$

(iii) If $n_f=\sum_{i=1}^{\rho}\tau_i$, $n_\varepsilon=\sum_{i=1}^{p}\varepsilon_i+p$, $\tilde{n}=n_f+n_\varepsilon$, $\tilde{\ell}=p$, then for every $\mathcal{S}(A,B)\in\mathcal{E}_K(A,B)$, $A\in\mathbb{R}^{\tilde{n}\times\tilde{n}}$, $B\in\mathbb{R}^{\tilde{n}\times\tilde{\ell}}$ and $\tilde{n}$ is the minimal dimension of $\mathcal{S}(A,B)$ systems for which parts (i) and (ii) hold true.

(iv) Let $\mathcal{S}(A,B)\in\mathcal{E}_K(A,B)$ and N be a left annihilator of B. The triple $(N,NA;\mathbb{R}^n)$ is a m.r.i.r. of the prime pair (F,G) and $\mathcal{S}_{i.r.}(A,B;\mathbb{R}^n)$ is a m.r.f.i.r. of (F,G). □

The Kronecker orbit $\mathcal{E}_K(A,B)$, derived under input, state coordinate transformations and state feedback, is known as the Brunovsky orbit of the pair (A,B); $\mathcal{E}_K(A,B)$ will thus be called the <u>natural Brunovsky orbit</u> (n.B.o.) of the prime pair (F,G). The orbit $\mathcal{E}_K(A,B)$ of (F,G) is characterised by a canonical element, the generalised Brunovsky canonical form [Kar. & MacB.-1] which is defined next.

<u>Remark (7.11)</u>: The natural Brunovsky orbit $\mathcal{E}_K(A,B)$ of the prime pair (A,B) is characterised by a canonical pair (A_K,B_K), the generalised

Brunovsky canonical form, which is defined by

$$A_K = \mathrm{diag}\{A_\varepsilon, A_f\},\ B_K = \begin{bmatrix} B_\varepsilon \\ 0 \end{bmatrix} \tag{7.59}$$

where $A_\varepsilon, B_\varepsilon, A_f$ are defined by $I(F,G)$ as shown by Corollary (7.15). □

Note that the n.B.o. is also parametrised as in Remark (7.10) if we set $K=0$, $C_K=0$, $G=0$. The results given above for the r.i.r. of $(F,G;\mathbb{R}^n)$, or (F,G), may be readily extended to the more general case of r.i.r. of $(F,G;V)$, $V \in \mathbb{R}^n$, as shown by the following result.

<u>Proposition (7.15)</u>: Let $(F,G) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$ and let $(N,NA;P)$ be a m.r.i.r. of $(F,G;\mathbb{R}^n)$, or (F,G). For every subspace $V \in \mathbb{R}^n$, $\dim V = d$, there exists a subspace $V' \subset P$, $\dim V' = d$, such that the triple $(N,NA;V')$ is a m.r.i.r. of $(F,G;V)$. □

The above results establish the existence of m.r.i.r., as well as m.r.f.i.r., and provide the means for their construction. Clearly, for a given pair (F,G), or a triple $(F,G;V)$, such realizations are not uniquely defined; in fact an orbit of m.r.i.r. is defined for every given pair (F,G), or a triple $(F,G;V)$. Nonminimal r.i.r. may be readily defined by an appropriate augmentation of the blocks in the triple (A,B,C) which corresponds to a m.r.f.i.r. of (F,G). For the case of prime pencils the Kronecker order of (F,G) will be referred to as the <u>Brunovsky order</u> of the prime pair. For the case of entirely right singular pencils we have:

<u>Remark (7.12)</u>: Let $(F,G) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$ be an entirely right singular pair ($sF-G$ is entirely right singular). The n.B.o. of (F,G), $E_K(A,B)$, is made up from controllable pairs (A,B). Furthermore, $E_K(A,B)$ is characterised by a canonical element, the Brunovsky canonical form which

is defined by (A_K,B_K), where $A_K=A_\varepsilon$, $B_K=B_\varepsilon$ and $A_\varepsilon,B_\varepsilon$ are defined by $\mathcal{I}(F,G)$ as shown by Corollary (7.15). □

The essence of the m.r.i.r. of a triple $(F,G;\mathcal{V})$ is that it reduces the study of properties of $\mathcal{V}$ with respect to (F,G) (expressed by the algebraic, geometric structure of $(F,G)/\mathcal{V}$) to an equivalent problem defined on a minimal triple $(N,NA;\tilde{\mathcal{V}})$ for which the pair (N,NA) is prime, or entirely right singular. The Kronecker order of a m.r.i.r. is defined by the set of strict equivalence invariants $\mathcal{I}(F,G)$ of (F,G) as shown by Theorem (7.10). Note that if $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $\tilde{n}\geq n$ and equality holds if and only if (F,G) is prime. Finally, note that using the notion of a m.r.i.r. we may provide an asymptotic characterisation for e-(F,G)-i.e., e-(G,F)-i.s. and e.c.-(F,G)-i.s. In fact, if $\mathcal{V}$ is a simple $\{\infty\}$-(F,G)-e.d.s., then we may always construct a m.r.i.r., $(N,NA;\tilde{\mathcal{V}})$, of $(F,G;\mathcal{V})$, where (N,NA) is entirely right singular. Given that $sN-NA$ may be considered as the restricted input state pencil of a controllable pair (A,B), then $\tilde{\mathcal{V}}$ is an infinite e.d. subspace of the system $\mathcal{S}(A,B)$ and the asymptotic results of Jaffe and Karcanias, 1981, [Jaf. & Kar.-1] apply for the characterisation of $\tilde{\mathcal{V}}$ and thus of $\mathcal{V}$. We should point out however that the basis tools used in this characterisation, i.e. the sequences of generalised B-spaces $\{\mathcal{V}_s\}$ with B-value s, are subspaces of $\mathbb{R}^{\tilde{n}}$ and not of $\mathbb{R}^n$. The notion of r.i.r. and r.f.i.r. of a triple $(F,G;\mathcal{V})$ plays an important role in the study of the solution space of autonomous generalised differential systems $\mathcal{S}(F,G)$ as will be shown next.

7.5 The solutions of the differential system $\mathcal{S}(F,G)$ and the dynamics of the fundamental invariant subspaces

7.5.1 Introduction

The aim of this section is the study of properties of the solution space of the differential systems

$$\mathcal{S}(F,G):\ F\dot{\underline{x}}(t)=G\underline{x}(t)\Leftrightarrow(pF-G)\underline{x}(t)=\underline{0} \qquad (7.60a)$$

$$\hat{\mathcal{S}}(F,G):\ F\hat{\underline{x}}(t)=G\dot{\hat{\underline{x}}}(t)\Leftrightarrow(F-pG)\hat{\underline{x}}(t)=\underline{0} \qquad (7.60b)$$

associated with the pair $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$. Note that $p\triangleq d/dt$ is the derivative operator and $\underline{x}(\cdot),\hat{\underline{x}}(\cdot):\mathbb{R}_+\to\mathbb{R}^n$, are called respectively the <u>generalised state</u> (g.s.), <u>dual generalised state</u> (d.g.s.) of the pair (F,G); the space $\mathbb{R}^n$ will be referred to as the <u>state domain</u> of the pair (F,G). The differential systems $\mathcal{S}(F,G),\hat{\mathcal{S}}(F,G)$ are assumed to be excited by arbitrary initial conditions $\underline{x}(0^-),\hat{\underline{x}}(0^-)\in\mathbb{R}^n$. The assumption of arbitrariness of initial conditions is essential to our analysis, since by making it we may also treat the important case of inconsistent initial conditions [Verg.-1],[Cam.-1]. An implicit assumption in the study of solutions of general differential systems describing dynamical systems is that "the dynamical systems have existed for a period prior to t=0 also". Under this condition the initial values at $t=0^-$, which determine the system solution, are themselves constrained to satisfy the system [Ros.-1]. The resulting constraints then guarantee that no impulsive behaviour occurs at t=0, since in effect any impulses are moved back in time to the instant of formation of the system. By considering unconstraint initial values, we may treat the case of systems formed at t=0, for which the initial values at $t=0^-$ satisfy the differential equations of a different dynamical system (see the behaviour of electric networks at "short-circuit" conditions etc. [Verg.-1]).

Differential systems of the type

$$D(p)\underline{\xi}(t) = N(p)\underline{u}(t) \qquad t\geq 0 \qquad (7.61)$$

where $D(p)\in\mathbb{R}[p]^{\nu\times\nu}$, $N(p)\in\mathbb{R}^{\nu\times\ell}$, D(p) nonsingular, have been studied by Vergese [Verg.-1] and Callier & Desoer [Cal. & Des.-1] etc. The emphasis in [Cal. & Des.-1] is on consistent initial conditions, whereas in [Verg.-1] both consistent and inconsistent initial conditions are examined. Differential systems of the type

$$A\dot{\underline{x}}(t)+B\underline{x}(t) = \underline{f}(t) \qquad t\geq 0 \tag{7.62}$$

where $A,B \in \mathbb{R}^{\nu\times\nu}$, $\underline{f}(t)$ input vector, have been studied by Campbell [Cam.-1], Vergese [Verg.-1], Wilkinson [Wil.-2] etc.; in [Cam.-1] and [Verg.-1] the case of both consistent and inconsistent initial conditions is considered. The differential systems $S(F,G)$,$\hat{S}(F,G)$ arise in the study of linear systems and in a sense are more general than those defined in (7.61) and (7.62), since the pencil $sF-\hat{s}G$ is allowed to be singular. It will be shown that in general systems of the $S(F,G)$ type do not define dynamical systems, but they are related to dynamical systems in a specific way. The distinguishing feature of $S(F,G)$ differential systems, when they are compared to those described by (7.61),(7.62), are the nonuniqueness of the solutions and the redundancy in the description of the state vector; the behaviour to consistent and inconsistent initial conditions is common to those described by (7.61),(7.62).

The type of solution of the differential system $S(F,G)$ ($\hat{S}(F,G)$) depends on the type of functions over which the differential system is studied. In the following we shall consider the class of C^{∞} (infinitely differentiable functions) and the class D_B' of Bohl distributions. Note that the class of time functions, whose Laplace transform is a strictly proper rational vector, is precisely the class of C^{∞}-functions on $t>0$, that are sums of exponential polynomials in t of the type $\sum_{k=0}^{m}(a_k t^k)\exp(\lambda t)$, $\forall t>0^-$. By definition, a Bohl distribution u may be expressed as $u=u_{imp}+u_{reg}$, where $u_{imp}=\sum_{i=0}^{k}\alpha_i\delta^{(i)}(t)$ and u_{reg} is a C^{∞}-function; thus $u\in D_B'$, if and only if its Laplace transform is a rational function. Clearly, D_B' contains as a subset the C^{∞} class, which in this context are referred to as regular Bohl distributions, or Bohl functions; note that $u\in D_B'$ is called impulsive if $u_{reg}=0$. The importance of the extension of the class C^{∞} to the distributions to the D_B' distributions is that it allows the handling of the difficult situations associated with nonconsistent initial conditions. Finally note that the class of scalar Bohl distributions D_B' forms a field with convolution as multiplication.

In the following we shall consider the differential system

$$S(F,G): F\dot{\underline{x}}(t)=G\underline{x}(t),\ (F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n} \tag{7.63}$$

and the results may be readily translated for the $\hat{S}(F,G)$, by using the properties of the e.d. type of duality.

Definition (7.10): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $t_o\in\mathbb{R}$ and $\underline{c}\in\mathbb{R}^n$.

(i) The vector $\underline{c}$ is said to be a consistent initial vector (c.i.v.) of $S(F,G)$ at t_o, if (7.63) possesses at least one solution from the space of C^∞-functions, or the space of D_B'-distributions, with $\underline{x}(t_o)=\underline{c}$. If the solution is from C^∞, D_B' spaces, then $\underline{c}$ will be referred to more specifically as C^∞-consistent initial vector (C^∞-c.i.v.), D_B'-consistent initial vector (D_B'-c.i.v.) respectively. The vector $\underline{c}$ will be called an inconsistent initial vector (i.i.v.) of $S(F,G)$, if (7.63) has no solution from the space of D_B'-distributions, when $\underline{x}(t_o)=c$. The space of i.i.v. will be denoted by C^* and shall be referred to as the redundancy space of $S(F,G)$. The space of c.i.v. will be denoted by C and shall be referred to as the initial space of $S(F,G)$.

(ii) A vector set $\underline{c}\in C$ will be called regular, if the initial value problem (7.63) with $\underline{x}(t_o)=\underline{c}$ has a unique solution; otherwise, i.e. the solution is not uniquely defined, it will be called nonregular. The subsets of C of regular, nonregular c.i.v. will be denoted by $C_{r.}, C_{n.r.}$ respectively.

(iii) The differential system $S(F,G)$ will be called regular if $C^*=\emptyset$ and every c.i.v. $\underline{c}$ is regular, or equivalently, when $C=C_r=\mathbb{R}^n$; otherwise, i.e. $C^*\neq\emptyset$, or $C^*=\emptyset$ and $C_{n.r.}\neq\emptyset$, the differential system $S(F,G)$ will be called singular. □

The above definition extends the notion of consistent initial condition defined by Campbell [Cam.-2], since the notion of consistency is extended to the case where a solution exists in the space of D_B' distributions. In the literature, [Verg.-1],[Cam.-1,2], the term "inconsistent initial condition" is used for $S(F,G)$ systems, described by regular pencils, for which the initial value problem has no solution from the C^∞ space (although a solution from the D_B' space exists). Here, the term "inconsistent initial vector" is reserved for the general case, where the initial value problem has no solution even from larger space of D_B' distributions. The terms "regular", "singular", used for the differential system $S(F,G)$ will be shown to be in exact correspondence to the characterisation of $S(F,G)$ based on the nature (regular,singular) of the associated matrix pencil.

We define as a normal state trajectory (n.s.t.) of $S(F,G)$ any C^∞-solution $\underline{x}(\cdot)$: $(0^-,\infty)\rightarrow\mathbb{R}^n$ of $(pF-G)\underline{x}(t)=\underline{0}$, $t\geq0$; hence by definition $\underline{x}(0^-)=\underline{x}(0^+)$. Similarly, we define as a distributional state trajectory (d.s.t.) of $S(F,G)$ any D_B'-solution $\underline{x}(\cdot)$: $(0^-,\infty)\rightarrow\mathbb{R}^n$ of $(pF-G)\underline{x}(t)=0$, $t\geq0$. In general, for the latter case, $\underline{x}(0^-)\neq\underline{x}(0^+)$. By an impulsive state trajectory (i.s.t.) we mean a d.s.t., $\underline{x}(t)\in D_B'^n$, which is expressed as $\underline{x}(t)=\sum_{i=0}^{k}\underline{x}_i\delta^i(t)$, $\underline{x}_i\in\mathbb{R}^n$. The set of solutions of $S(F,G)$ which correspond to all $\underline{c}\in S_r$ will be denoted by X_r and will be called the regular solution space of $S(F,G)$. Let us denote by C_r^c,C_r^d the subsets of C_r which yield C^∞,D_B' but not C^∞ solutions respectively, and by X_r^c,X_r^d the subsets of X_r which correspond to C_r^c,C_r^d respectively. X_r^c,X_r^d will be called the C^∞-regular solution space, D_B'-regular solution space correspondingly and $\mathcal{X}_r=\mathcal{X}_r^c\cup\mathcal{X}_r^d$. The set of solutions of $S(F,G)$, which correspond to all $\underline{c}\in\mathcal{C}_{n.r.}$ will be denoted by $\mathcal{X}_{n.r.}$ and will be called the nonregular solution space of $S(F,G)$. The set of solutions which correspond to all $\underline{c}\in\mathcal{C}$ will be called the solution space of $S(F,G)$ and shall be denoted by $\mathcal{X}$; clearly, $\mathcal{X}=\mathcal{X}_r\cup\mathcal{X}_{n.r.}$ Note that $\mathcal{X}_{n.r.}$ may have elements from the D_B' space and/or from the C^∞ space.

In this section we shall concentrate on the study of the $\mathcal{X}$ solution space and the $\mathcal{C}$ initial space, which generates $\mathcal{X}$. Of special interest is the characterisation of the corresponding subsets of $\mathcal{X},\mathcal{C}$ defined before. Such a study is intimately related to the nature of the set $\mathcal{I}(F,G)$ of strict equivalence invariants of the pencil sF-G associated with $S(F,G)$. It will be shown that each distinct type of invariant characterises a distinct dynamic property of $S(F,G)$ and leads to a characterisation of the subsets of $\mathcal{X},\mathcal{C}$. The results derived, provide the means for the dynamic characterisation of the various types of invariant subspaces of the domain of (F,G). In the following, unless it is specifically mentioned, it will be assumed that $\mathcal{I}(F,G)=\{\mathcal{I}_c(F,G);\mathcal{I}_r(F,G);\mathcal{D}_o(F,G);\mathcal{D}_\infty(F,G);\mathcal{D}_\alpha(F,G)\}$ where the individual types of invariants have the general expression given in (7.1); a differential system with such a structure of invariants will be referred to as a general differential system.

7.5.2 The differential system $S(F,G)$ and strict equivalence transformations

We start off by examining the effects of strict equivalence transformations on the spaces $\mathcal{X}$, $\mathcal{C}$ and $\mathcal{C}^*$. We examine first the nature of $\mathcal{X}$ and then the effect of left, right, left-right strict equivalence transformations on the spaces $\mathcal{X},\mathcal{C}^*$ of the differential system $S(F,G)$.

Proposition (7.16): The solution space $\mathcal{X}$ of a general $\mathcal{S}(F,G)$ differential system is an $\mathbb{R}$-vector space. □

The proof of this result is rather obvious. In fact, if $\underline{x}_1(t),\underline{x}_2(t)$ are solutions of (7.63), then by observing that the real matrices and the differential operator are linear it follows that $\alpha_1\underline{x}_1(t)+\alpha_2\underline{x}_2(t)\in\mathcal{X}$, $\forall\alpha_1,\alpha_2\in\mathbb{R}$. Note, that the above arguments hold true for any $\underline{x}_1(t),\underline{x}_2(t)$ stimulated by any initial conditions in $\mathcal{C}$.

Proposition (7.17): Let $(F,G),(F',G')\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $R\in\mathbb{R}^{m\times m}$, $|R|\neq 0$, and let $F'=RF$, $G'=RG$. Consider the differential systems

$$(pF-G)\underline{x}(t)=\underline{0},\ t\geq 0,\quad (pF'-G')\underline{x}'(t)=\underline{0},\ t\geq 0 \tag{7.64}$$

with solution spaces $\mathcal{X},\mathcal{X}'$ and redundancy spaces $\mathcal{C}^*,\mathcal{C}^{*'}$ respectively. Then, $\mathcal{X}=\mathcal{X}'$ and $\mathcal{C}^*=\mathcal{C}^{*'}$. □

Once more the proof is rather straightforward and it is omitted. For the case of right strict equivalence we have.

Proposition (7.18): Let $(F,G),(\tilde{F},\tilde{G})\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$, $Q\in\mathbb{R}^{n\times n}$, $|Q|\neq 0$ and let $\tilde{F}=FQ$, $\tilde{G}=GQ$. Consider the differential systems

$$(pF-G)\underline{x}(t)=\underline{0},\ t\geq 0,\quad (p\tilde{F}-\tilde{G})\tilde{\underline{x}}(t)=\underline{0},\ t\geq 0 \tag{7.65}$$

with solution spaces $\mathcal{X},\tilde{\mathcal{X}}$ and redundancy spaces $\mathcal{C}^*,\tilde{\mathcal{C}}^*$ respectively. The maps f: $\tilde{\mathcal{X}}\to\mathcal{X}$: $\tilde{\underline{x}}(t)\in\tilde{\mathcal{X}}\to f(\tilde{\underline{x}}(t))\triangleq Q\tilde{\underline{x}}(t)=\underline{x}(t)\in\mathcal{X}$ and f*: $\tilde{\mathcal{C}}^*\to\mathcal{C}^*$: $\tilde{\underline{c}}\in\mathcal{C}^*\to f^*(\tilde{\underline{c}})=Q\tilde{\underline{c}}=\underline{c}\in\mathcal{C}^*$ are linear bijections (isomorphisms) from $\tilde{\mathcal{X}}$ onto $\mathcal{X}$ and $\tilde{\mathcal{C}}^*$ onto $\mathcal{C}^*$ correspondingly. □

The proof is obvious and it is omitted. By combining the last two propositions we have.

Theorem (7.11): Let (F,G), $(F'',G'') \in \mathbb{R}^{m\mathrm{x}n} \mathrm{x} \mathbb{R}^{m\times n}$, $R \in \mathbb{R}^{m\mathrm{x}m}$, $Q \in \mathbb{R}^{n\mathrm{x}n}$, $|R|,|Q| \neq 0$ and let $F''=RFQ$, $G''=RGQ$. Consider the differential systems

$$(pF-G)\underline{x}(t)=\underline{0}, \quad t\geq 0 \qquad (pF''-G'')\underline{x}''(t)=\underline{0}, \quad t\geq 0 \tag{7.66}$$

with solution spaces X, X'' and redundancy spaces C^*, $C^{*''}$ respectively. The maps $f: X''\to X: \underline{x}''(t) \in X'' \to f(\underline{x}''(t)) \triangleq Q\underline{x}''(t) = \underline{x}(t) \in X$ and $f^*: C^{*''}\to C^*: \underline{c}'' \in C^{*''} \to f(\underline{c}'') \triangleq Q\underline{c}''= \underline{c} \in C^*$ are linear bijections (isomorphisms) from X'' onto X and $C^{*''}$ onto C^* correspondingly.

□

Theorem (7.11) implies that if $\underline{x}''(t)$ is a solution of $S(F'',G'')$, then $Q\underline{x}''(t)=\underline{x}(t)$ is a solution of $S(F,G)$ and vice versa. Thus, strict equivalence transformations may be used for simplifying the description of $S(F,G)$ and then studying the properties of the solution space on a simpler desctiption. In the following the Kronecker canonical form, (F_k,G_k), of the pair (F,G) will be used and the results will be finally translated back to the original description. Thus, let (R,Q), $R \in \mathbb{R}^{m\mathrm{x}m}$, $Q \in \mathbb{R}^{n\mathrm{x}n}$, $|R|,|Q| \neq 0$ such that

$$(F,G) \underset{(R^{-1},Q^{-1})}{\overset{(R,Q)}{\rightleftarrows}} (RFQ,RGQ) = (F_k,G_k) \tag{7.67a}$$

where (F_k,G_k) have the following canonical description

$$F_k = \text{block-diag.}\{F_\zeta;F_\varepsilon;F_f;F_\infty\}, \quad G_k = \text{block-diag.}\{G_\zeta;G_\varepsilon;G_f;G_\infty\} \tag{7.67b}$$

The pairs (F_ζ,G_ζ), $(F_\varepsilon,G_\varepsilon)$, (F_f,G_f), (F_∞,G_∞) are associated with the canonical blocks of the Kronecker form characterising the sets of r.m.i., c.m.i., f.e.d., i.e.d. respectively. Clearly, the pair (R,Q) is not uniquely defined. The reduction to Kronecker canonical form of (F,G) implies the following reduction on $S(F,G)$

$$S(F,G): F\dot{\underline{x}}=G\underline{x} \underset{(R^{-1},Q^{-1})}{\overset{(R,Q)}{\rightleftarrows}} S(F_k,G_k): F_k\dot{\underline{x}}'= G_k\underline{x}', \quad \underline{x}=Q\underline{x}' \tag{7.67c}$$

The g.s. $\underline{x}'(t)$ of $S(F_k,G_k)$ will be called the Kronecker generalised state. The spaces X_k, C_k^*, C_k associated with $S(F_k,G_k)$ will be referred to as the Kronecker solution space, Kronecker redundancy space, Kronecker initial space respectively of the differential system $S(F,G)$ and they characrerise the strict equivalence orbit of $S(F,G)$ systems and not the invividual $S(F,G)$.

The precise relationship of X_k, C_k^*, C_k to the spaces X, C^*, C of $S(F,G)$ is defined by the isomorphism introduced by the matrix Q. If we now partition $\underline{x}'(t)$ according to the partitioning of (F_k,G_k) i.e. $\underline{x}'(t)=[\underline{x}_\zeta(t)^t;\underline{x}_\varepsilon(t)^t;\underline{x}_f(t)^t;\underline{x}_\infty(t)^t]^t$, then $S(F_k,G_k)$ is equivalent to the following set of subsystems

$$S(F_k,G_k):F\dot{\underline{x}}'(t)=G\underline{x}'(t)\Leftrightarrow\left[\begin{array}{ll} S(F_\zeta,G_\zeta):F_\zeta\dot{\underline{x}}_\zeta(t)=G_\zeta\underline{x}_\zeta(t) & (7.68)\\ S(F_\varepsilon,G_\varepsilon):F_\varepsilon\dot{\underline{x}}_\varepsilon(t)=G_\varepsilon\underline{x}_\varepsilon(t) & (7.69)\\ S(F_f,G_f):F_f\dot{\underline{x}}_f(t)=G_f\underline{x}_f(t) & (7.70)\\ S(F_\infty,G_\infty):F_\infty\dot{\underline{x}}_\infty(t)=G_\infty\underline{x}_\infty(t) & (7.71)\end{array}\right.$$

By studying the properties of the subsystems $S(F_\zeta,G_\zeta)$, $S(F_\varepsilon,G_\varepsilon)$, $S(F_f,G_f)$, $S(F_\infty,G_\infty)$ the results may be transferred to $S(F_k,G_k)$ and thus finally to the original description $S(F,G)$. In the following we shall examine the case of pencils characterised by only one type of invariants and possibly zero r.m.i.; the results are then used to provide a dynamic characterisation of the subspaces V of the domain of a general pair (F,G) for which $(F,G)/V$ is characterised by only one type of invariants. The basis for the dynamic characterisation of subspaces is the nature of the solutions of the general system $S(F,G)$ which are restricted in the given subspace V for initial conditions taken from V; such solutions are defined by the solutions of the V-restricted differential system $S(FV,GV)$. Finally, the solution space properties of a general system $S(F,G)$ are examined and the dynamic properties of the various types of invariant subspaces is given.

7.5.3. Entirely left singular differential systems

Let $(F,G)\in \mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ and $S(F,G)$ be the associated differential system. $S(F,G)$ will be called <u>entirely left singular</u> (e.l.s.) if (F,G), or the associated pencil $sF-G$ is entirely left singular. For e.l.s. pairs (F,G), $I(F,G)= I_r(F,G)=\{\zeta_1=\ldots=\zeta_{g'}=0,\ 0<\zeta_{g'+1}\leq\ldots\leq\zeta_t\}$ and thus the Kronecker form $(F_k,G_k)=(F_\zeta,G_\zeta)$. Thus, (7.67c) may be written as

$$S(F,G):F\underline{\dot{x}}=G\underline{x}\ \underset{(R^{-1},Q^{-1})}{\overset{(R,Q)}{\rightleftarrows}}\ S(F_\zeta,G_\zeta):F_\zeta\underline{\dot{x}}_\zeta=G_\zeta\underline{x}_\zeta,\quad \underline{x}=Q\underline{x}_\zeta \tag{7.72a}$$

where

$$F_\zeta=\begin{bmatrix} & O_{g',n} & \\ L_{\zeta_{g'+1}} & & O \\ & \ddots & \\ O & & L_{\zeta_t}\end{bmatrix},\qquad G_\zeta=\begin{bmatrix} & O_{g',n} & \\ \hat{L}_{\zeta_{g'+1}} & & O \\ & \ddots & \\ O & & \hat{L}_{\zeta_t}\end{bmatrix} \tag{7.72b}$$

and

$$L_\zeta=\begin{bmatrix} I_\zeta \\ \hline \underline{0}^t\end{bmatrix}\in\mathbb{R}^{(\zeta+1)x\zeta},\qquad \hat{L}_\zeta=\begin{bmatrix} \underline{0}^t \\ \hline I_\zeta\end{bmatrix}\in\mathbb{R}^{(\zeta+1)x\zeta} \tag{7.72c}$$

For a general differential system having an e.l.s. part we have the following result [Kar. & Hay. -2].

<u>Proposition (7.19)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$, $S(F,G)$ be the associated differential system and let $n_r=\sum_{i=1}^{t}\zeta_i$ be the number associated with the set of r.m.i., $I_r(F,G)$ of (F,G).

(i) The system $S(F,G)$ is e.l.s. if and only if $n=n_r$. Then, $m=n_r+t$ (t=number of r.m.i.).

(ii) If $n_r=n$, i.e. $S(F,G)$ is e.l.s., then the only solution of $S(F,G)$ (over the D'_β distributions space) is $\underline{x}(t)=\underline{0}$, $t\geq 0$.

(iii) If $0<n_r<n$, i.e. $\mathcal{S}(F,G)$ is a general differential system, then there is a number n_r of independent linear relations among the coordinates of the generalised state vector $\underline{x}(t)$ of $\mathcal{S}(F,G)$.

<u>Proof</u>

(i) The first part follows by inspection of the Kronecker form.

(ii) By (7.72) and if $\underline{x}_\zeta=[x^t_{\zeta_{g'+1}};\dots;\underline{x}^t_{\zeta_t}]$, then for the non zero blocks we have:

$$L_{\zeta_j}\dot{x}_{\zeta_j} = \hat{L}_{\zeta_j}\underline{x}_{\zeta_j}, \qquad j=g'+1,\dots,t \tag{7.73a}$$

Because of the (7.72c) structure of $(L_{\zeta_j},\hat{L}_{\zeta_j})$ blocks, it is readily shown that the only solution of (7.73a) for $\forall$ $j=g'+1,\dots,t$ is the zero solution. Since $\underline{x}_\zeta=\underline{0}$ and Q invertible we have that $\underline{x}(t)=\underline{0}$, $t\geq 0$.

(iii) If $0<n_r<n$, then $\mathcal{S}(F,G)$ is a general differential system. Let (R,Q) be a pair of strict equivalence transformations thae reduces (F,G) to the Kronecker form (F_k,G_k). We may write $(F_k,G_k)=(\text{block diag.}\{F_\zeta,\tilde{F}\}, \text{block diag.}\{G_\zeta,\tilde{G}\})$, where $(\tilde{F},\tilde{G})$ corresponds to all other invariants apart from r.m.i. By partitioning $\underline{x}_k$ accordingly, i.e. $\underline{x}_k=[\underline{x}^t_\zeta,\underline{\tilde{x}}^t]^t$, $\mathcal{S}(F_k,G_k)$ is reduced to the following two subsystems:

$$F_\zeta\dot{\underline{x}}_\zeta = G_\zeta\underline{x}_\zeta \quad \text{and} \quad \tilde{F}\dot{\underline{\tilde{x}}} = \tilde{G}\underline{\tilde{x}} \tag{7.73b}$$

By part (ii) $\underline{x}_\zeta(t)=0$, $t\geq 0$ is the only solution of $\mathcal{S}(F_\zeta,G_\zeta)$ and thus

$$\underline{x}_k = \begin{bmatrix} \underline{0} \\ \hline \underline{\tilde{x}} \end{bmatrix} = \hat{Q}\underline{x} = \begin{bmatrix} Q_1 \\ \hline Q_2 \end{bmatrix}\underline{x}, \qquad \hat{Q}=Q^{-1}$$

from which

$$Q_1\underline{x} = \underline{0}, \quad Q_1\in\mathbb{R}^{n_r\mathrm{x}n} \quad \text{and} \quad \underline{\tilde{x}} = Q_2\underline{x} \tag{7.73c}$$

the first of the above conditions expresses the existence of n_r linear relationships (Q_1 has full rank) among the coordinates of $\underline{x}$.

□

From the above result it follows that if $\mathcal{S}(F,G)$ is e.l.s. the $\mathcal{C}=\{0\}$ and the redundancy space $\mathcal{C}^*=\mathbb{R}^n-\{0\}$. For a general differential system with $n_r>0$ the above result demonstrates that the coordinates of generalised state vector are not independent, but they satisfy n_r linear relations. The index n_r has been previously defined as the left-(F,G)-order of $\mathbb{R}^n$, $C_\ell(F,G;\mathbb{R}^n)$; because it expresses a number of independent relationships amongst the coordinates of $\underline{x}(t)$ it will be also referred to as the <u>reduncancy index</u> of $\mathcal{S}(F,G)$. If $n_r=0$ (i.e. (F,G) has only zero r.m.i.) depence relationships of the type described above do not exist amongst the coordinates of $\underline{x}(t)$. The investigation of whether there exist other depence relationships, which of course is connected with the nature of the $\mathcal{C}^*$ space, cannot be answered unless we examine the solvability of the other subsystems (7.69), (7.70), (7.71); however we may state the following remark.

<u>Remark (7.13)</u>: The redundancy space $\mathcal{C}^*$ of a general differential system $\mathcal{S}(F,G)$ is a nonempty set if the left-(F,G)-order of $\mathbb{R}^n$, $C_\ell(F,G;\mathbb{R}^n)>0$. Furthermore, if the domain $\mathbb{R}^n$ is expressed as in (7.38a) i.e.

$$\mathbb{R}^n = \mathcal{V}_\zeta \oplus \mathcal{V}_\epsilon \oplus \mathcal{V}_o \oplus \mathcal{V}_\alpha \oplus \mathcal{V}_\infty \qquad (7.74)$$

where $\mathcal{I}_{\mathcal{V}_\zeta}=\{\mathcal{I}_r(F,G);\mathcal{I}_r^o\}$, then every $\underline{c}\in\mathcal{V}_\zeta$, $\underline{c}\neq\underline{0}$, is also a vector in the $\mathcal{C}^*$ space of $\mathcal{S}(F,G)$.

□

The above results enable us to give the following important dynamic property of the subspaces of the domain of a general pair (F,G).

<u>Preposition (7.20)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$, $\mathcal{V}\subset\mathbb{R}^n$ be a subspace and let $\dim\mathcal{V}=d$. If the left-(F,G)-order of $\mathcal{V}$, $C_\ell(F,G;\mathcal{V})>0$, then there exist initial vectors $\underline{x}(0)\in\mathcal{V}$ such that trajectories of $\mathcal{S}(F,G)$ starting from $\underline{x}(0)$, leave the subspace $\mathcal{V}$.

Proof

To show that for some $\underline{x}(0) \in \mathcal{V}$ there exist trajectories that leave $\mathcal{V}$ we have to show that the restricted differential system $\mathcal{S}(FV,GV)$, V is a basis matrix of $\mathcal{V}$, has no solution for some initial vectors $\underline{v}(0)$; equivalently, we have to show that the $\mathcal{C}^*$ of (FV,GV) is nonempty. By Remark (7.13) and $C_\ell(F,G;\mathcal{V})>0$ the result is established.

$\square$

7.5.4. Entirely right singular differential systems

Let $(F,G) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$ and $\mathcal{S}(F,G)$ be the associated differential system. $\mathcal{S}(F,G)$ will be called entirely right singular (e.r.s.) if (F,G), or the associated pensil sF-G is entirely right singular. An extension of the notion of e.r.s. differential systems may be introduced by allowing in $\mathcal{I}(F,G)$ also a set of zero r.m.i.; differential systems of this type will be referred to as extended entirely right singular (e.e.r.s.).

Remark (7.14): Let $\mathcal{S}(F,G):F\dot{\underline{x}}(t)=G\underline{x}(t)$ be an e.e.r.s. differential system. There always exist transformations $R \in \mathbb{R}^{m\times m}$, $|R|\neq 0$, such that

$$\mathcal{S}(F,G) \underset{(R^{-1},I)}{\overset{(R,I)}{\rightleftarrows}} \mathcal{S}(F',G'):\begin{bmatrix} 0 \\ --- \\ \tilde{F} \end{bmatrix} \dot{\underline{x}}(t) = \begin{bmatrix} 0 \\ --- \\ \tilde{G} \end{bmatrix} \underline{x}(t) \tag{7.75}$$

where $\mathcal{S}(\tilde{F},\tilde{G}):\tilde{F}\dot{\underline{x}}(t):\tilde{G}\underline{x}(t)$ is e.r.s. Clearly, if $\mathcal{X}$, $\tilde{\mathcal{X}}$ are the solution spaces of $\mathcal{S}(F,G)$, $\mathcal{S}(\tilde{F},\tilde{G})$ respectively, then $\mathcal{X}=\tilde{\mathcal{X}}$.

$\square$

From the above remark it is clear that the study of e.e.r.s. is identical to the study of equivalent e.r.s. systems and this in the following we shall consider the case of e.r.s. systems. For e.r.s. pairs (F,G), $\mathcal{I}(F,G)=\mathcal{I}_c(F,G)=\{\varepsilon_1=\ldots=\varepsilon_g=0<\varepsilon_{g+1}\leq\ldots\leq\varepsilon_p\}$ and thus the Kronecker form $(F_k,G_k)=(F_\varepsilon,G_\varepsilon)$. Thus (7.67c) may be written as

$$S(F,G):F\underline{\dot{x}}(t)=G\underline{x}(t) \underset{(R^{-1},Q^{-1})}{\overset{(R,Q)}{\rightleftarrows}} S(F_\varepsilon,G_\varepsilon):F_\varepsilon\underline{\dot{x}}_\varepsilon(t)=G_\varepsilon\underline{x}_\varepsilon(t),\ \ \underline{x}=Q\underline{x}_\varepsilon \tag{7.76a}$$

where

$$F_\varepsilon = \left[\begin{array}{c|c|c} & L_{\varepsilon_{g+1}} & 0 \\ \cline{2-2} 0_{m,g} & & \ddots \\ \cline{3-3} & 0 & L_{\varepsilon_p} \end{array}\right], \qquad G_\varepsilon = \left[\begin{array}{c|c|c} & \hat{L}_{\varepsilon_{g+1}} & \\ \cline{2-2} 0_{m,g} & & \ddots \\ \cline{3-3} & 0 & \hat{L}_{\varepsilon_p} \end{array}\right] \tag{7.76b}$$

$$L_\varepsilon = \left[I_\varepsilon \,\vdots\, \underline{0}\right] \in \mathbb{R}^{\varepsilon x(\varepsilon+1)}, \qquad \hat{L}_\varepsilon = \left[\underline{0} \,\vdots\, I_\varepsilon\right] \in \mathbb{R}^{\varepsilon x(\varepsilon+1)} \tag{7.76c}$$

For an e.r.s. differential system we may state the following result [Kar. & Hay. -2].

Proposition (7.21): Let $(F,G) \in \mathbb{R}^{mxn} x \mathbb{R}^{mxn}$ be an e.r.s. pair, $S(F,G)$ be the associated differential system and let $I(F,G)=I_r(F,G)=\{\varepsilon_i,\ i\in\underset{\sim}{p}\}$.

(i) For all $\underline{x}(0^-) \in \mathbb{R}^n$, there exists a family of solutions $\Sigma(\underline{x}(0^-))$ of $S(F,G)$ which is nontrivial (contains at least one non zero elelment).

(ii) There are p independent linear combinations of the coordinates of the g.s.v. $\underline{x}(t)$ which may be arbitrarily assigned for $\forall\ \underline{x}(0^-) \in \mathbb{R}^n$.

Proof

We seek solutions over the space of generalised functions which have Laplace transforms. We start from the canonical description (7.76); by taking Laplace transforms [Doe. -1] with $\underline{x}_\varepsilon(0^-)$ (the transformed to the new frame initial condition) we have

$$(sF_\varepsilon - G_\varepsilon)\underline{\hat{x}}_\varepsilon(s) = F_\varepsilon\underline{x}_\varepsilon(0^-) \tag{7.77a}$$

where $\underline{\hat{x}}_\varepsilon(s) = L\{\underline{x}_\varepsilon(t)\}$. If we now write

$$s\underline{L}_{\varepsilon}-\hat{L}_{\varepsilon} = \left[\begin{array}{ccccc|c} s & -1 & 0 & \dots & 0 & 0 \\ 0 & s & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & s & -1 \end{array}\right] = [T_{\varepsilon}(s), -\underline{e}_{\varepsilon}] \in \mathbb{R}^{\varepsilon x(\varepsilon+1)}[s] \tag{7.77b}$$

and partition $\underline{\hat{x}}_{\varepsilon}(s)$ as

$$\underline{\hat{x}}_{\varepsilon}(s) = [\underbrace{\underline{\hat{x}}_{g}(s)^{t}}_{g}; \underbrace{\underline{\hat{x}}_{\varepsilon_{g+1}}(s)^{t}}_{\varepsilon_{g+1}+1}; \dots; \underbrace{\underline{\hat{x}}_{\varepsilon_{p}}(s)^{t}}_{\varepsilon_{p}+1}]^{t} \tag{7.77c}$$

$\underline{\hat{x}}_{\varepsilon_j}(s) = [x_1^j(s), \dots, x_{\varepsilon_j}^j(s), x_{\varepsilon_j+1}^j(s)]^t = [\underline{\hat{x}}'^{\,t}_{\varepsilon_j}(t); x^j_{\varepsilon_{j+1}}(s)]^t$, j=g+1,...,p then (7.77a) yields the equivalent system of equations

$$0_{m,g} \cdot \underline{\hat{x}}_{g}(s) = 0_{m,g} \cdot \underline{x}_{g}(0^{-}) \tag{7.78a}$$

$$\left[\begin{array}{ccc} T_{\varepsilon_{g+1}}(s) & & 0 \\ & \ddots & \\ 0 & & T_{\varepsilon_p}(s) \end{array}\right] \left[\begin{array}{c} \underline{\hat{x}}'_{\varepsilon_{g+1}}(s) \\ \vdots \\ \underline{\hat{x}}'_{\varepsilon_p}(s) \end{array}\right] = \left[\begin{array}{ccc} \underline{e}_{\varepsilon_{g+1}} & & 0 \\ & \ddots & \\ 0 & & \underline{e}_{\varepsilon_p} \end{array}\right] \left[\begin{array}{c} x^{g+1}_{\varepsilon_{g+1}+1}(s) \\ \vdots \\ x^{p}_{\varepsilon_p+1}(s) \end{array}\right] + \tilde{F}_{\varepsilon}\tilde{x}_{\varepsilon}(0^{-}) \tag{7.78b}$$

where $\underline{\tilde{x}}_{\varepsilon}(0^{-})$ is the part of $\underline{x}_{\varepsilon}(0^{-})$ that corresponds to the non zero blocks of F_{ε}; $\tilde{F}_{\varepsilon}$ is the part of F_{ε} that excludes the zero block. Clearly, for all $\underline{x}_g(0^{-})$ and all admissible functions (7.78a) is trivially satisfied. Since diag$\{T_{\varepsilon_j}(s)$, j=g+1,...,p$\}$ has full rank over $\mathbb{R}(s)$, a family of solutions of (7.78b) exists over the space of admissible functions for arbitrary $\underline{\tilde{x}}_{\varepsilon}(0^{-})$ and arbitrary choice of $\{\underline{x}^{g+1}_{\varepsilon_{g+1}}(s), \dots, \underline{x}^{p}_{\varepsilon_p+1}(s)\}$ (from the space of admissible functions). By transforming the result buck to the original frame (S(F,G) system) the result is established.

$\square$

The above result establishes the existence of a family of solutions $\Sigma(\underline{x}(0^{-}))$ for $\forall$ $\underline{x}(0^{-}) \in \mathbb{R}^n$ if S(F,G) is e.r.s., or e.e.r.s. The monuniqueness of the solution, clearly implies that the differential system $S(F,G)$ does not define a dynamical system; however, it is related to forced dynamical systems. This relationship will be investigated next, and the elements of the $\Sigma(\underline{x}(0^-))$ family will be parametrised in terms of external inputs of an equivalent system. Of crucial importance in this investigation is the notion of m.r.f.i.r. of the pair (F,G) or of the triple $(F,G;\mathbb{R}^n)$. It is worth pointing out that the family of solutions $\Sigma(\underline{x}(0^-))$ is defined over the space D_L^n, where D_L is the space of distributions with a Laplace transform [Doe. -1]; this family will be denoted by $\Sigma_D(\underline{x}(0^-))$, in short. A subset of D_L are the Bohl distributions D_B' (distributions with a rational Laplace transform); the subset of $\Sigma_D(\underline{x}(0^-))$, defined over $D_B'^n$ will be denoted by $\Sigma_B(\underline{x}(0^-))$. For the study of holdability properties of trajectories in a subspace, it will be shown that the $\Sigma_B(\underline{x}(0^-))$ family is sufficient for our investigation.

<u>Remark (7.15)</u>: For every $\underline{x}(0^-)\in\mathbb{R}^n$, the families $\Sigma_D(\underline{x}(0^-))$, $\Sigma_B(\underline{x}(0^-))$ of the e.r.s. differential system $S(F,G)$ are $\mathbb{R}$-vector spaces. □

By Corollary (7.16) we have:

<u>Proposition (7.22)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be an e.r.s. pair, $I(F,G)=I_r(F,G)=\{\varepsilon_i,\ i\in\underset{\sim}{p}\}$ and let $E_K(A,B)$ be the n.B.o. of (F,G).

(i) The $E_K(A,B)$ orbit is controllable and it is characterised by the set $\{\sigma_i:\sigma_i=\varepsilon_i+1,\ i\in\underset{\sim}{p}\}$ of controllability indices.

(ii) Let $S(A,B)\in E_K(A,B)$ where

$$S(A,B):\ \dot{\underline{\tilde{x}}}(t)=A\underline{\tilde{x}}(t)+B\underline{\tilde{u}}(t),\quad \underline{\tilde{x}}\in\mathbb{R}^n,\ \underline{\tilde{u}}\in\mathbb{R}^p \tag{7.79a}$$

$sN-NA\in\mathbb{R}^{mxn}[s]$ be the corresponding restricted input-state pencil and let $sF-G=R(sN-NA)Q$, $R\in\mathbb{R}^{mxm}$, $Q\in\mathbb{R}^{nxn}$, $|R|,\ |Q|\neq0$.

If $\Sigma_D^{A,B}(\underline{\tilde{x}}(0^-)$ is the family of solutions of $S(A,B)$, that correspond to hte given $\underline{\tilde{x}}(0^-)$ and all $\underline{\tilde{u}}(t)\in D_L^p$ and $\Sigma_D(\underline{x}(0^-))$ the family of solutions of

$\mathcal{S}(F,G)$ which correspond to $\underline{x}(0^-)=Q\underline{\tilde{x}}(0^-)$, then $\Sigma_D^{A,B}(\underline{\tilde{x}}(0^-))$ and $\Sigma_D(\underline{x}(0^-))$ are isomorphic. Furthermore, the family $\Sigma_D(\underline{x}(0^-))$ is given by

$$\underline{x}(t) = Q\{e^{At}\underline{\tilde{x}}(0^-) + \int_{0^-}^{t} e^{A(t-\tau)}B\underline{\tilde{u}}(\tau)d\tau\} \quad \forall\ \underline{\tilde{u}}(t)\in D_L^p \tag{7.79b}$$

Proof

Part (i) follows from Corollary (7.16). For any $\mathcal{S}(A,B)\in\mathcal{E}_K(A,B)$ we have that

$$\mathcal{S}(A,B):\ \dot{\underline{\tilde{x}}}=A\underline{\tilde{x}}+B\underline{\tilde{u}} \Leftrightarrow N\dot{\underline{\tilde{x}}}=NA\underline{\tilde{x}},\quad \tilde{u}=B^+(\dot{\underline{\tilde{x}}}-A\underline{\tilde{x}}) \tag{7.80}$$

where N,B^+ are left annihilators, inverse of B. Furthermore, the pencil sN-NA and sF-G are strict equivalent and thus, their solution spaces are isomorphic i.e. every solution of $\mathcal{S}(F,G)$ defines a solution of $\mathcal{S}(N,NA)$ and vice versa. The differentiability, in the distributional sense of the elements of $\Sigma_D(\underline{x}(0^-))$ is essential for the definition of $\underline{\tilde{u}}(t)$ in the equivalence defined by (7.80).

□

The above result demonstrates that the solution space $\mathcal{X}$ of the e.r.s. system $\mathcal{S}(F,G)$ is isomorphic to the solution space of a proper m.r.f.r. described by the $\mathcal{S}(A,B)$: $\dot{\underline{x}}=A\underline{x}+B\underline{u}$ forced linear system with (A,B) controllable; thus we may state

Remark (7.16): The properties of the solution space $\Sigma_D(x(0^-))$ of an e.r.s. differential system $\mathcal{S}(F,G)$ are completely described by the properties of the solution space of a controllable forced linear system $\mathcal{S}(A,B)$ which corresponds to a given initial condition $\underline{x}(0^-)$ and any control input $\underline{u}(t)\in D_L^p$. Furthermore, every solution $\underline{x}(t)\in\Sigma_D(x(0^-))$ may be parametrised by a $\underline{u}(t)\in D_L^p$ defined by

$$\underline{u}(t) = B^+\{\dot{\underline{x}}(t)-A\underline{x}(t)\} \tag{7.81}$$

□

From the above results, it is obvious that the controllability properties of the standard linear theory may be transferred to the case of autonomous differential systems. This topic will be examined next. Before we proceed, it is worth pointing out that the $\Sigma_B(x(0^-))$ family of solutions contains as a subset those solutions which are defined from the space of C^∞ functions (functions with a strictly proper rational Laplace transform). This subfamily will be denoted by $\Sigma_c(\underline{x}(0^-))$.

Proposition (7.23): Let $(F,G) \in \mathbb{R}^{mxn}$ be an e.r.s. pair, $S(F,G)$ the associated differential system; $\underline{c}_1, \underline{c}_2 \in \mathbb{R}^n$ and let $\Sigma_B(\underline{c}_1)$, $\Sigma_c(\underline{c}_1)$ be the families of solutions of $S(F,G)$ which correspond to $x(0^-)=\underline{c}_1$.

(i) For $\forall$ $(\underline{c}_1,\underline{c}_2)$ pair, there exists a distributional trajectory $x(t) \in \Sigma_B(\underline{c}_1)$ such that $\underline{x}(0^+)=\underline{c}_2$.

(ii) For $\forall$ $(\underline{c}_1,\underline{c}_2)$ pair, there exists a normal state trajectory $\underline{x}(t) \in \Sigma_c(\underline{c}_1)$ such that $\underline{x}(t_o^+)=\underline{c}_2$ for some $t_o \geq 0$.

□

This result follows immediately be Remark (7.16) and the standard results on reachability of linear systems (see [Won. -1],[Kai. -1]). Part (i) expresses the property that any two points in the space $\mathbb{R}^n$ may be connected by a distributional trajectory in "zero time" [Kar. & Kour. -2],[Kai. -1], while part (ii) expresses the standard notion of reachability [Kal. -1]. The above property motivates the following definition.

Definition (7.11): Let $(F,G) \in \mathbb{R}^{mxn} x \mathbb{R}^{mxn}$ be a general pair, $S(F,G)$ the associated differential system, and C the initial space of $S(F,G)$. For a $\underline{c}_1 \in C$, we shall denote by $\Sigma_B(\underline{c}_1)$, $\Sigma_c(\underline{c}_1)$ the families D_B'-distributional, C^∞-trajectories of $S(F,G)$ which are excited by $\underline{x}(0^-)=\underline{c}_1$. We define:

(i) A pair $(\underline{c}_1,\underline{c}_2) \in C x \mathbb{R}^n$ will be said to be D_B'-reachable, C^∞-reachable, if there exists $\underline{x}(t) \in \Sigma_B(\underline{c}_1)$, $\underline{x}'(t) \in \Sigma_c(c_1)$ respectively and $t_o \geq 0$ such that $(\underline{x}(0^-)=\underline{c}_1, \underline{x}(t_o^+)=\underline{c}_2)$, $(\underline{x}'(0^-)=\underline{c}_1, \underline{x}'(t_o^+)=\underline{c}_2)$ correspondingly.

(ii) The set of vectors $\underline{c}\in\mathbb{R}^n$ for which $(\underline{0},\underline{c})$ is D_B'-reachable, C^∞-reachable will be denoted by $\mathcal{R}_B$, $\mathcal{R}_C$ respectively; $\mathcal{R}_B$, $\mathcal{R}_C$ will be called the $\underline{D_B'\text{-reachable set}}$, $\underline{C^\infty\text{-reachable set}}$ of $\mathcal{S}(F,G)$. The set $\mathcal{R}_B^*$, $\mathcal{R}_C^*$ which are complementary of $\mathcal{R}_B$, $\mathcal{R}_C$ respectively with respect to $\mathbb{R}^n$ will be referred to as the $\underline{D_B'\text{-nonreachable-}}$, $\underline{C^\infty\text{-nonreachable-set}}$ of $\mathcal{S}(F,G)$ correspondingly.

(iii) If every pair $(\underline{c}_1,\underline{c}_2)\in C\times\mathbb{R}^n$ is D_B'-reachable, C^∞-reachable then the system (F,G) will be called $\underline{D_B'\text{-reachable}}$, $\underline{C^\infty\text{-reachable}}$ respectively. □

Because of the linearity we may state the following:

Remark (7.17): $\mathcal{S}(F,G)$ is $\underline{D_B'\text{-reachable}}$, $\underline{C^\infty\text{-reachable}}$ if every pair $(\underline{0},\underline{c})$, $\underline{c}\in\mathbb{R}^n$ is D_B'-reachable, C^∞-reachable respectively. □

Clearly, this remark simplifies the study of reachability. For the sets $\mathcal{R}_B$, $\mathcal{R}_C$, linearity implies the following property.

Remark (7.18): The sets $\mathcal{R}_B$, $\mathcal{R}_C$ of a general system $\mathcal{S}(F,G)$ are $\mathbb{R}$-linear vector spaces. Furthermore, $\mathcal{S}(F,G)$ is D_B'-reachable, C^∞-reachable if and only if $\mathcal{R}_B=\mathbb{R}^n$, $\mathcal{R}_C=\mathbb{R}^n$ respectively. □

By Proposition (7.23) and Remark (7.14) we have:

Remark (7.19): If $\mathcal{S}(F,G)$ ia an e.e.r.s. differential system, then it is D_B'-reachable and C^∞-reachable. □

The converse of the latter result will be studied in the last section. Our attention is focused next on subspaces of the domain of a general pair (F,G). Motivated by the standard results of the geometric theory [Won. -1], [Bas. & Mar. -1], [Will. -1] we give the following definition for $\mathcal{S}(F,G)$ systems.

Definition (7.12): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ be a general pair, $\mathcal{S}(F,G)$ the associated differential system, $\Sigma_B(\underline{c})$, $\Sigma_C(\underline{c})$ the D_B'-distributional-, C^∞-families of trajectories for a $\underline{x}(0^-)=\underline{c}\in\mathbb{R}^n$ and let $V\in\mathbb{R}^n$ be a subspace.

(i) The subspace V will be called a C^∞-holding subspace (C^∞-h.s.), if for $\forall\ \underline{x}(0^-)=\underline{c}\in V$, $\exists\ \underline{x}(t)\in\Sigma_C(\underline{c})$ such that $\underline{x}(t)\in V$, $\forall\ t\in\mathbb{R}^+$.

(ii) The subspace V will be called a D'_B-distributionally holding subspace (D'_B-h.s.), if for $\forall\ \underline{x}(0^-)=\underline{c}\in V$, $\exists\ x(t)\in\Sigma_B(\underline{c})$ such that $\underline{x}(t)\in V$, $\forall\ t\in\mathbb{R}^+$.

(iii) The subspace V will be called a C^∞-reachability subspace (C^∞-r.s.), if for $\forall\ \underline{c}_1,\underline{c}_2\in V$, $\exists\ t_o\geq 0$ and $\underline{x}(t)\in\Sigma_C(\underline{c}_1)$ such that $\underline{x}(0^-)=\underline{c}_1$, $\underline{x}(t_o^+)=\underline{c}_2$ and $\underline{x}(t)\in V$, $\forall\ t\in\mathbb{R}^+$.

(iv) The subspace V will be called a D'_B-distributionally reachability subspace (D'_B-r.s.), if for $\forall\ \underline{c}_1,\underline{c}_2\in V$, $\exists\ t_o\geq 0$ and $x(t)\in\Sigma_B(\underline{c}_1)$, such that $\underline{x}(0^-)=\underline{c}_1$, $x(t_o^+)=\underline{c}_2$ and $\underline{x}(t)\in V$, $\forall\ t\in\mathbb{R}^+$. $\square$

We will denote by $H_C(F,G)$, $H_B(F,G)$, $K_C(F,G)$, $K_B(F,G)$ the families of all C^∞-h.s., D'_B-h.s., C^∞-r.s., D'_B-r.s., respectively defined on a pair (F,G). From the definition, the following properties may be readily verified.

Remark (7.20): The families of subspaces H_C, H_B, K_C, K_B defined on a general differential system $S(F,G)$ satisfy the properties.

$$K_C\subset H_C,\quad K_B\subset H_B$$

$\square$

To show the above properties take $\underline{c}_2=\underline{0}$ in the definition of C^∞-r.s., D'_B-r.s. The families of subspaces defined above are extensions os the standard notions of the geometric theory to the case of $S(F,G)$ differential systems. The links between the standard dynamic notions and those defined here, will be explored in the last section by using the notion of m.r.f.i.r. of $(F,G;V)$. The properties of subspaces mentioned above may be checked by investigating the properties of solution space of differential systems associated with V.

Definition (7.13): Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ be a general pair, $S(F,G):F\dot{\underline{x}}(t)=G\underline{x}(t)$, be the associated differential system, $V\subset\mathbb{R}^n$ be a subspace and V_α

basis matrix of $\mathcal{V}$. We define

$$\mathcal{S}(FV,GV):FV\underline{\dot{v}}(t)=GV\underline{v}(t),\quad \underline{x}(t)=V\underline{v}(t) \tag{7.82}$$

as the <u>$\mathcal{V}$-restricted differential system</u> of $\mathcal{S}(F,G)$. The initial, redundancy space of $\mathcal{S}(FV,GV)$ shall be denoted by $C_{\mathcal{V}}$, $C^{*}_{\mathcal{V}}$ respectively. The subset of consisting of all $\underline{c}\in C_{\mathcal{V}}$ for which there exists at least one C^{∞}-trajectory will be denoted by $C^{c}_{\mathcal{V}}$. Clearly, if $\dim\mathcal{V}=d$, then $C_{\mathcal{V}}$, $C^{*}_{\mathcal{V}}$, $C^{c}_{\mathcal{V}}$ are subsets of $\mathbb{R}^{d}$. All other definitions given for $\mathcal{S}(F,G)$ carry over in a natural way the case of $\mathcal{S}(FV,GV)$.

□

By definitions (7.12), (7.13) and Proposition (7.20) we have:

<u>Proposition (7.24)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ be a general pair, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace, $\dim\mathcal{V}=d$, $C_{\ell}(F,G;\mathcal{V})$ the left-(F,G)-order of $\mathcal{V}$ and let $\mathcal{S}(FV,GV)$ be the $\mathcal{V}$-restricted differential system of $\mathcal{S}(F,G)$.

(i) Necessary and sufficient conditions for $\mathcal{V}$ to be a D'_{B}-h.s., is that $C_{\mathcal{V}}=\mathbb{R}^{d}$.

(ii) Necessary and sufficient condition for $\mathcal{V}$ to be a C^{∞}-h.s., is that $C^{c}_{\mathcal{V}}$.

(iii) Necessary and sufficient condition for $\mathcal{V}$ to be a C^{∞}-r.s. are thar $C^{c}_{\mathcal{V}}=\mathbb{R}^{d}$ and $\mathcal{S}(FV,GV)$ is C^{∞}-reachable.

(iv) Necessary and sufficient conditions for $\mathcal{V}$ to be a D'_{B}-r.s. are that $C_{\mathcal{V}}=\mathbb{R}^{d}$ and $\mathcal{S}(FV,GV)$ is D'_{B}-reachable.

(v) Necessary condition for $\mathcal{V}$ to be C^{∞}-r.s., D'_{B}-h.s., C^{∞}-r.s., or D'_{B}-r.s. is that $C_{\ell}(F,G;\mathcal{V})=0$.

□

This result will be extensively used throughout the rest of this chapter. We close this section by investigating the properties of $\mathcal{I}_{c}$-(F,G)-c.m.i.s. in the context of the above definitions.

<u>Theorem (7.12)</u>: Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ be a general pair and let $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace. If $\mathcal{V}$ is an $\mathcal{I}_{c}$-(F,G)-c.m.i.s. then,

(i) $\mathcal{V}$ is a C^{∞}-h.s., as well as a D'_{B}-h.s.

(ii) $\mathcal{V}$ is a C^{∞}-r.s., as well as a D_B'-r.s. □

The proof of the result follows from the fact that $S(FV,GV)$ is an e.e.r.s. differential system; then by Remark (7.14), Propositions (7.23) and (7.24) and Definition (7.13) the results follows. The families of C^{∞}-, D_B'-solutions defined for $\forall\ \underline{x}(0^-)=\underline{c}\in\mathcal{V}$ will be denoted by $\Sigma_C^{\mathcal{V}}(\underline{x}(0^-))$, $\Sigma_B^{\mathcal{V}}(\underline{x}(0^-))$ and they may be parametrised as it is indicated by Proposition (7.22) and Remark (7.16).

7.5.5. Regular differential systems

Let $(F,G)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}$ and $S(F,G)$ be the associated differential system. $S(F,G)$ will be called extended entirely right regular (e.e.r.r.), if (F,G), or the associated pensil $sF-G$ is right nonsingular and the left-(F,G)-order of $\mathbb{R}^n$ $C_{\ell}(F,G;\mathbb{R}^n)=0$. Clearly, $S(F,G)$ is e.e.r.r. if $sF-G$ is characterised by e.d. and possible zero r.m.i.; for e.e.r.r. systems we have that $m\geq n$. If $m=n$ and $S(F,G)$ is e.e.r.r., then it will be called entirely regular (e.r.) For e.r. differential systems we have that $sF-G$ is a regular pencil.

Remark (7.21): Let $S(F,G)$ be an e.e.r.r. differential system. There always exist transformations $R\in\mathbb{R}^{m\times m}$, $|R|\neq 0$, such that

$$S(F,G):F\dot{\underline{x}}(t)=G\underline{x}(t) \underset{(R^{-1},I)}{\overset{(R,I)}{\rightleftarrows}} S(F',G'),\ \begin{bmatrix} 0 \\ --- \\ \tilde{F} \end{bmatrix}\dot{\underline{x}}(t)=\begin{bmatrix} 0 \\ --- \\ \tilde{G} \end{bmatrix}\underline{x}(t) \qquad (7.83)$$

where $S(\tilde{F},\tilde{G}):\tilde{F}\dot{\underline{x}}(t)=\tilde{G}\underline{x}(t)$ is e.r. Clearly, if X, $\tilde{X}$ are the solution spaces of $S(F,G)$, $S(\tilde{F},\tilde{G})$ respectively, then $X=\tilde{X}$. □

From the above remark it follows that the study of e.e.r.r. systems is identical to that of e.r. systems and thus, in the following we shall concentrate on e.r. $S(F,G)$ systems. Let $(F,G)\in\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times n}$, (F,G) regular, $\mathcal{D}_{\infty}(F,G)\triangleq\{\hat{s}^{q_i},\ i\in\underset{\sim}{\mu}\}$ the set of i.e.d., $\mathcal{D}_f(F,G)\triangleq\{(s-\lambda_i)^{d_i},\ i\in\underset{\sim}{\tau}\}$ the set of all finite e.d. of (F,G); we shall denote by $\Theta_{\infty}=\sum_{i=1}^{\mu} q_i$ and $\Theta_f=\sum_{i=1}^{\tau} d_i$ and thus

$\Theta_f=n-\Theta_\infty$. There exists pairs (R,Q) of strict equivalence transformations (real, or complex) that reduce (F,G) to the Weierstrass form (real, or complex) (F_w,G_w) and thus

$$\mathcal{S}(F,G):F\dot{\underline{x}}(t)=G\underline{x}(t)\ \underset{(R^{-1},Q^{-1})}{\overset{(R,Q)}{\rightleftarrows}}\ \mathcal{S}(F_w,G_w):F_w\dot{\underline{x}}_w(t)=G_w\underline{x}_w(t) \tag{7.84a}$$

where

$$F_w=\text{diag}\{H_\infty;\ I_{\theta_f}\}\ , \qquad G_w=\text{diag}\{I_{\theta_\infty};\ C_f\} \tag{7.84b}$$

$H_\infty=\text{diag}\{H_{q_i}\ ,\ i\in\underline{\mu}\}\in\mathbb{R}^{\theta_\infty x\theta_\infty}$, H_{q_i} the $q_i x q_i$ standard nilpotent block and $C_f\in\mathbb{R}^{\theta_f x\theta_f}$ is the Jordan form, or real canonical form associated with the set $\mathcal{D}_f(F,G)$. By partitioning the state vector $\underline{x}_w$ according to the (7.84b) partitioning, i.e. $\underline{x}_w=[\underline{x}_\infty^t;\ \underline{x}_f^t]^t$, (7.84a) is reduced to the following system of differential systems

$$\mathcal{S}(F_w,G_w)\iff\left[\begin{array}{ll} H_\infty\dot{\underline{x}}_\infty(t)=\underline{x}_\infty(t) & (7.85a)\\ \dot{\underline{x}}_f(t)=C_f\underline{x}_f(t) & (7.85b)\end{array}\right.$$

The numbers θ_∞, θ_f defined above, characterise the dimensions of the subsystems in (7.85a), (7.85b) and shall be referred to as the <u>infinite order</u>, <u>finite order</u> of the e.r. $\mathcal{S}(F,G)$, as well as of any e.e.r.r. system. For a regular initial space $\mathcal{C}_r$ we have defined the subsets $\mathcal{C}_r^c$, $\mathcal{C}_r^d$ which correspond to C^∞-regular solutions, D_B'-regular but not C^∞-solutions respectively. The subset of $\mathcal{C}_r$ which yield purely impulsive solutions will be denoted by $\mathcal{C}_r^\delta$ and the corresponding solution space will be denoted by $\mathcal{X}_r^\delta$; $\mathcal{X}_r^\delta$ will be referred to as the <u>purely impulsive regular solution space</u>. Clearly, $\mathcal{C}_r^\delta\subset\mathcal{C}_r^d\subset\mathcal{C}_r$ and $\mathcal{X}_r^\delta\subset\mathcal{X}_r^d\subset\mathcal{X}_r$. For an e.r. (F,G) system we have:

<u>Theorem (7.13)</u>: Let $(F,G)\in\mathbb{R}^{nxn}x\mathbb{R}^{nxn}$ be an e.r. pair, θ_∞, θ_f be the infinite, finite order of (F,G), $\mathcal{S}(F,G)$ the associated differential system and let $\mathcal{C}$, $\mathcal{X}$ be the initial space, solution space of $\mathcal{S}(F,G)$ respectively.

(i) $\forall\ \underline{c}\in\mathbb{R}^n$ is a consistent, regular initial vector and thus $\mathcal{C}=\mathcal{C}_r=\mathbb{R}^n$ and $\mathcal{X}=\mathcal{X}_r$.

(ii) The subsets $\mathcal{C}_r^c$, $\mathcal{C}_r^\delta$ are $\mathbb{R}$-linear vector spaces, $\dim\mathcal{C}_r^c=\theta_f$, $\dim\mathcal{C}_r^\delta=\theta_\infty$ and $\mathcal{C}_r^d=\mathbb{R}^n-\mathcal{C}_r^c$.

(iii) $\mathcal{X}=\mathcal{X}_r$ is an $\mathbb{R}$-linear vector space and $\dim\mathcal{X}=n$. Furthermore the subsets $\mathcal{X}_r^c$, $\mathcal{X}_r^\delta$ of $\mathcal{X}_r$ are $\mathbb{R}$-linear vector spaces; $\dim\mathcal{X}_r^c=\theta_f$, $\dim\mathcal{X}^\delta=\theta_\infty-\mu$, ($\mu=$ number of i.e.d.), $\mathcal{X}=\mathcal{X}_r\oplus\mathcal{X}^\delta$ and $\dim\mathcal{X}=n-\mu$.

Proof

(i) Solving $\mathcal{S}(F,G):F\dot{\underline{x}}(t)=G\underline{x}(t)$ by the Laplace transform method, for arbitrary initial values of $\underline{x}(t)$ at $t=0^-$ [Doe. -1],[Kai. -1] we get the equation:

$$(sF-G)\hat{\underline{x}}(s)=F\underline{x}(0^-) \tag{7.86a}$$

where $\hat{\underline{x}}(s)=\mathcal{L}_-\{\underline{x}(t)\}$ is the Laplace transform from 0^- of $\underline{x}(t)$. Since $(sF-G)$ is regular (invertible over $\mathbb{R}(s)$) we have that

$$\hat{\underline{x}}(s)=(sF-G)^{-1}F\underline{x}(0^-) \tag{7.86b}$$

Equation (7.86b) gives the solution $\underline{x}(t)=\mathcal{L}_-^{-1}\{\underline{x}(s)\}$ of $\mathcal{S}(F,G)$ for all $\underline{x}(0^-)\in\mathbb{R}^n$. Clearly, for $\forall\ \underline{x}(0^-)=\underline{c}\in\mathbb{R}^n$ a solution exists and it is uniquely defined; furthermore, since $(sF-G)^{-1}F\in\mathbb{R}^{n\times n}(s)$, $\underline{x}(t)\in\mathcal{D}_B'^n$ (belongs the space of Bohl distributions). Thus, $\mathcal{C}=\mathcal{C}_r=\mathbb{R}^n$ and $\mathcal{X}$ is regular, i.e. $\mathcal{X}=\mathcal{X}_r$.

(ii) Let (R,Q) be a pair such that $(F,G)=(\hat{R}F_w\hat{Q},\ \hat{R}G_w\hat{Q})$, $\hat{R}=R^{-1}$, $\hat{Q}=Q^{-1}$ and $\underline{x}(\cdot)=Q\underline{x}_w(\cdot)$. Then,

$$\hat{\underline{x}}(s)=Q\{(sF_w-G_w)^{-1}F_w\}\underline{x}_w(0^-),\qquad \underline{x}(0^-)=Q\underline{x}_w(0^-) \tag{7.86c}$$

Using (7.84b), the above expression yields

$$\hat{\underline{x}}(s)=Q\begin{bmatrix}(sH_\infty-I_{\theta_\infty})^{-1}H_\infty & 0\\ 0 & (sI_{\theta_f}-C_f)^{-1}\end{bmatrix}\underline{x}_w(0^-) \tag{7.86d}$$

$\underline{x}_w(0^-)$ may now be considered as an arbitrary initial vector of $\mathbb{R}^n$ (since $\underline{x}(0^-)=Q\underline{x}_w(0^-)$, $Q\in\mathbb{R}^{n\times n}$, $|Q|\neq 0$). By partitioning $\underline{x}_w(0^-)=\underline{c}$ as $\underline{c}=[\underline{c}_\infty^t, \underline{c}_f^t]$ and $Q=[Q_\infty, Q_f]$ according to the (7.84b) partitioning, eqn. (7.86) yields

$$\hat{\underline{x}}(s)=Q_\infty(sH_\infty-I_{\theta_\infty})^{-1}H_\infty\underline{c}_\infty+Q_f(sI_{\theta_f}-C_f)^{-1}\underline{c}_f \tag{7.87}$$

Note that $(sH_\infty-I_{\theta_\infty})^{-1}H_\infty\in\mathbb{R}^{\theta_\infty x\theta_\infty}[s]$ polynomial, since $sH_\infty-I_{\theta_\infty}$ is $\mathbb{R}[s]$-unimodular and $(sI_{\theta_f}-C_f)^{-1}\in\mathbb{R}^{\theta_f x\theta_f}(s)$ strictly proper. Thus for $\forall\ \underline{x}(0^-)\in$ sp$\{Q_f\}$, $\hat{\underline{x}}(s)\in\mathbb{R}^n(s)$ strictly proper and for $\forall\ \underline{x}(0^-)\in$sp$\{Q_\infty\}$, $\hat{\underline{x}}(s)\in\mathbb{R}^n[s]$ polynomial; furthermore, if $\underline{x}(0^-)$ has non zero projection on both sp$\{Q_\infty\}$, sp$\{Q_f\}$ subspaces, the $\hat{\underline{x}}(s)\in\mathbb{R}^n(s)$ is a general rational vector with non zero polynomial and non zero strictly proper part. Clearly $C_r^c=$sp$\{Q_f\}$ and $C_r^\delta=$sp$\{Q_\infty\}$ are $\mathbb{R}$-vector spaces and since Q_∞, Q_f have full rank we have that dim$C_r^\delta=$rank$(Q_\infty)=\theta_\infty$, dim$C_r^c=$rank$(Q_f)=\theta_f$. For all $\underline{c}\in\mathbb{R}^n$ such that $\underline{c}\notin$sp$\{Q_f\}$ the solution has always an impulsive part and thus $C_r^d=\mathbb{R}^n-C_r^c$.

(iii) Equation (7.87) clearly demonstrates that $\mathcal{X}_r$, $\mathcal{X}_r^\delta$ and $\mathcal{X}_r^c$ are $\mathbb{R}$-linear vector spaces. If we define $\hat{\Psi}(s)=[\hat{\Psi}_f(s), \hat{\Psi}_\infty(s)]$, where

$$\Psi_f(s)=Q_f(sI_{\theta_f}-C_f)^{-1}\in\mathbb{R}^{nx\theta_f}(s),\ \hat{\Psi}_\infty(s)=Q_\infty(sH_\infty-I_{\theta_\infty})^{-1}H_\infty\in\mathbb{R}^{nx\theta_\infty}[s] \tag{7.88}$$

then, rank$_{\mathbb{R}(s)}\hat{\Psi}_f(s)=\theta_f$, and it is readily shown that rank$_{\mathbb{R}(s)}\hat{\Psi}_\infty(s)=\theta_\infty-\mu$. Every column in $\hat{\Psi}_f(s)$, $\hat{\Psi}_\infty(s)$ defines a C^∞-solution, purely impulsive solution respectively. Every C^∞-solution may be uniquely expressed in terms of the solutions defined by the columns of $\hat{\Psi}_f(s)$ and every purely impulsive solution may be uniquely expressed in terms of the solutions corresponding to the non zero columns of $\hat{\Psi}_\infty(s)$. Thus the columns of $\hat{\Psi}_f(s)$ and the non zero columns in $\hat{\Psi}_\infty(s)$ are the Laplace transforms of solutions which form basis sets for $\mathcal{X}_r^c$, $\mathcal{X}_r^\delta$ respectively. Thus, dim$\mathcal{X}_r^c=\theta_f$, dim$\mathcal{X}_r^\delta=\theta_\infty-\mu$ and clearly (from (7.87)), $\mathcal{X}_r=\mathcal{X}_r^c\oplus\mathcal{X}_r^\delta$. Obviously, dim$\mathcal{X}_r=n-\mu$.

□

From the analysis of Chapter (4) (see Corollary (4.9)) we have that the matrix $\bar{Q}$ which is used in the reduction of (F,G) to the Weierstrass form (complex) may be expressed as

$$\bar{Q} = [\bar{Q}_{\infty},\bar{Q}_{\alpha_1},\dots,\bar{Q}_{\alpha_i},\dots,\bar{Q}_{\alpha_\mu}] \tag{7.89a}$$

where the columns of $\bar{Q}_{\alpha_i}$ form a complete prime set of chains of (F,G) at $s=\alpha_i\in\mathcal{R}$, ($\mathcal{R}=\{\infty,\alpha_1,\dots,\alpha_\mu\}$ is the root range of (F,G)).

Let $\bar{Q}_{\alpha_i}=[\dots;\underline{x}^{k}_{\alpha_i};\dots]$. For every pair of complex conjugate (α_i,α_i^*) values, we may substitute the pair of block $(\bar{Q}_{\alpha_i},\ \bar{Q}_{\alpha_i^*})$ in (7.89a) by a single real block

$$Q_{\alpha_i} = [\dots;\mathrm{Re}(\underline{x}^{k}_{\alpha_i});\mathrm{Im}(\underline{x}^{k}_{\alpha_i});\dots] \tag{7.89b}$$

The matrix defined by

$$Q = [Q_{\infty},Q_{\alpha_1},\dots,Q_{\alpha_i},\dots,Q_{\alpha_\nu}] \tag{7.89c}$$

where $Q_{\alpha_j}=\bar{Q}_{\alpha_j}$ if $\alpha_j\in\mathbb{R}\cup\{\infty\}$ may then be used for the reduction of (F,G) to the real spectral form. We shall denote by $\mathcal{M}^*_{\beta}$ the maximal generalised null space of (F,G) at $s=\beta\in\mathbb{R}$, $\beta\in\mathbb{R}\cup\{\infty\}$. For a pair of complex conjugate $\alpha,\alpha^*\in\mathbb{R}$ we define as the (α,α*) maximal generalised null space of (F,G) the space $\mathcal{M}^*_{\alpha,\alpha^*}$ defined by

$$\mathcal{M}^*_{\alpha,\alpha^*} = \mathrm{sp}_{\mathbb{R}}\{\mathrm{Re}(\underline{x}),\ \mathrm{I}_m(\underline{x}):\ \forall\ \underline{x}\in\mathcal{M}^*_{\alpha}\} \tag{7.89d}$$

Clearly, Q_{α} (as defined by (7.89b)) is a characteristic basis matrix of $\mathcal{M}^*_{\alpha,\alpha^*}$ with respect to the real spectral form. Similarly $Q_{\beta}=\bar{Q}_{\beta}$ (as defined in (7.89a)) is a characteristic basis matrix with respect to the real spectral form if $\beta\in\mathbb{R}\cup\{\infty\}$. The basis matrixes Q_{β}, $\beta\in\mathbb{R}\cup\{\infty\}$, Q_{α}, $\alpha\in\mathbb{C}-\mathbb{R}$, defined above will be referred to as β-, (α,α*)-real spectral characteristic basis matrix of the subspaces $\mathcal{M}^*_{\beta}$, $\mathcal{M}^*_{\alpha,\alpha^*}$ respectively. The matrix Q in (7.89c) may be expressed as

$$Q = [Q_\infty, Q_f], \quad Q_f = [\ldots, Q_{\alpha_i}, \ldots], \; \alpha_i \in \mathbb{R}-\{\infty\} \tag{7.90a}$$

and $\mathbb{R}^n$ may be written in a direct sum form as

$$\mathbb{R}^n = M^*_\infty \oplus M^*_f \tag{7.90b}$$

where M^*_f is the direct sum of all M^*_β, $M^*_{\alpha,\alpha*}$ spaces defined for $\beta \in R$, $\beta \in \mathbb{R}-\{\infty\}$ and $\alpha \in R, \alpha \in \mathbb{C}-\mathbb{R}$. M^*_∞, M^*_f will be called the *maximal generalised null space of (F,G) at $\{\infty\}$, $\mathbb{C}-\{\infty\}$* respectively; Q_∞, Q_f are *characteristic real spectral basis matrices* of M^*_∞, M^*_f correspondingly. The index of annihilation of (F,G) at $s=\infty$ (maximal order of i.e.d. of (F,G)) shall be denoted by q_∞.

Corollary (7.18): Let $(F,G) \in \mathbb{R}^{nxn} x \mathbb{R}^{nxn}$ be an e.r. pair, M^*_∞, M^*_f the maximal generalised null spaces of (F,G) at $\{\infty\}$, $\mathbb{C}-\{\infty\}$ and let Q_∞, Q_f be characteristic real spectral basis matrices for M^*_∞, M^*_f correspondingly. Then,

(i) $C^\delta_r = M^*_\infty$ and $C^c_r = M^*_f$

(ii) Define the matrices $\Psi_f(t), \Psi_\infty(t), \Psi(t) = [\Psi_f(t), \Psi_\infty(t)]$, $t \geq 0^-$, by

$$\Psi_f(t) = Q_f e^{C_f t}, \quad \Psi_\infty(t) = Q_\infty \sum_{i=0}^{q_\infty-2} \delta^{(i)}(t) H_\infty^{i+1} \tag{7.91}$$

(a) $\Psi_f(t)$ is a basis matrix for X^c_r. The non zero vectors in $\Psi_\infty(t)$, $\Psi(t)$ define basis vectors sets for X^δ_r, X respectively.

(b) If $\underline{x}(0^-) = \underline{x}_0 \in \mathbb{R}^n$, the general D'_B-solution of $S(F,G)$ may be written as

$$\underline{x}(t,x_0) = \Psi_f(t)\hat{Q}_f\underline{x}(0^-) - \Psi_\infty(t)\hat{Q}_\infty\underline{x}(0^-) \tag{7.92a}$$

where $\hat{Q}_f$, $\hat{Q}_\infty$ are $\theta_f xn$, $\theta_\infty xn$ matrices defined by

$$\hat{Q} = [Q_f, Q_\infty]^{-1} = \begin{bmatrix} \hat{Q}_f \\ \hline \hat{Q}_\infty \end{bmatrix} \tag{7.92b}$$

□

The analysis of the e.r. differential system and the study of its solutions has also appeared in the literature in many different forms. Campbell [Camp. -1,2] has examined the problems using the Drazin inverses and the solution for the homogeneous case is similar to that given here (eqn. (7.92)). A justification for the solution of $\mathcal{S}(F,G)$ was given by Cobb [Cob. -1]. Cobb's reasoning is important since it explains the distributional solution as the limit of solutions of systems with F invertible. We consider here the simpler system (unforced) and of the type

$$\mathcal{S}(H_\infty, I_{\theta_\infty}): H_\infty \underline{\dot{x}}(t) = \underline{x}(t), \qquad H_\infty \in \mathbb{R}^{nxn} \tag{7.93a}$$

For any $\underline{x}(0^-)=\underline{x}_o \in \mathbb{R}^n$, the formal solution of (7.93a) is

$$\underline{x}(t,\underline{x}_o) = -\sum_{i=0}^{q_\infty -2} \delta^{(i)}(t) H^{i+1} \underline{c} \tag{7.93b}$$

Following Cobb we define:

<u>Definition (7.14)</u>: The distributional vector $\underline{x}(t)$ is a <u>limiting solution</u> of $H_\infty \underline{\dot{x}}=\underline{x}$ with initial condition $\underline{x}_o$, if there exist sequences $\{H_j\} \subseteq \mathbb{R}^{nxn}$, $\{\underline{x}_{o_j}\} \subseteq \mathbb{R}^n$, with H_j invertible, such that $H_j \to H_\infty$, $\underline{x}_{o_j} \to \underline{x}_o$ and the solution of $H_j \underline{\dot{x}}_j = \underline{x}_j$, $\underline{x}_j(0^-)=\underline{x}_{o_j}$ converges to $\underline{x}$. ☐

<u>Theorem (7.14)[Cob. -1]</u>: For every $\underline{x}_o \in \mathbb{R}^n$, there exists a limiting solution of $H_\infty \underline{\dot{x}}=\underline{x}$ with initial condition $\underline{x}_o$. The limiting solution is unique and it is $\underline{x}(t,\underline{x}_o)$ in (7.93b).

☐

Francic [Fra. -1] has also examined the problem and he gives sufficient conditions for convergence in a distributional sence. We should point out that the work of Willems [Will. -1], Trentelman [Tren. -1] on distributionally controlled invariant subspaces (almost controlled invariant subspaces) and distributionally controllability subspaces (almost controllability subspaces) is intimately related to the present study; in fact the work of Karvanias [Kar. -1], Jaffe and Karcanias [Juf. & Kar. -1] has established the links between Willems characterisation and the matrix pencil characterisation of generalised invariant subspaces. It is through this equivalence and the notion of invariant forced realisation that the dynamic (distributional) characterisation in [Will. -1], [Tren. -1] may be transferred to our present study. Next, we give some results related to the dynamic characterisation of subspaces.

<u>Remark (7.22)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be a general pair, $\mathcal{V}\subset\mathbb{R}^{n}$, be a subspace $\dim\mathcal{V}=d$ and let $(N,NA;\tilde{\mathcal{V}})$, $\mathcal{S}_{i.r.}(A,B;\tilde{\mathcal{V}})$ be a p.m.r.i.r., p.m.r.f.i.r. of $(F,G;\mathcal{V})$ respectively. The solution space of $\mathcal{S}(F\mathcal{V},G\mathcal{V})$ is <u>equivalent</u> to the solution space of $\mathcal{S}_{i.r.}(A,B)$ in $\tilde{\mathcal{V}}$, i.e. every solution of $\mathcal{S}_{i.r.}(A,B)$ in $\mathcal{V}$ yields a solution of $\mathcal{S}(F\mathcal{V},G\mathcal{V})$ and vice versa.

□

This remark shows that the results in [Will. -1],[Tren. -1] may be interpreted to equivalent results on $\mathcal{S}(F,G)$ systems, as long as the restricted differential systems $\mathcal{S}(F\mathcal{V},G\mathcal{V})$, $\mathcal{S}(N\tilde{\mathcal{V}},NA\tilde{\mathcal{V}})$ are strict equivalent.

<u>Theorem (7.15)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be a general pair, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace and let $\Phi_{\mathcal{V}}$ be the root range of $(F,G)/\mathcal{V}$.

(i) If $\mathcal{V}$ is a Φ-(F,G)-e.d.s., then it is a D_B'-h.s.

(ii) If $\mathcal{V}$ is a Φ_∞,-(F,G)-e.d.s., (∞-proper, or f.e.d), then it is also a C^∞-h.s.

(iii) If $\mathcal{V}$ is a simple $\{\infty\}$-(F,G)-e.d.s., then it is a D_B'-h.s. and a D_B'-r.s.

□

Part (i) and (ii) follow immediately from Corollary (7.18). Note that if $\mathcal{V}$ is a simple $\{\infty\}$-(F,G)-e.d.s., then in a p.m.r.f.i.r. $\mathcal{S}_{i.r.}(A,B;\tilde{\mathcal{V}})$, the subspace $\tilde{\mathcal{V}}$ is a sliding subspace [Juf. & Kar. -1]. A rigorous proof for the distributional reachability property of sliding subspaces is given in [Tren. -1]. By combining part (iii) of Theorem (7.15) and Theorem (7.12) we have:

<u>Corollary (7.19)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be a general pair, $\mathcal{V}\subset\mathbb{R}^{n}$ be a subspace. If $(F,G)/\mathcal{V}$ is characterised by c.m.i., i.e.d. and possibly zero r.m.i., then $\mathcal{V}$ is a D_B'-h.s. and also a D_B'-r.s.

□

We close this section by noting that the general solution $\underline{x}(t,\underline{x}_0)$ in (7.92a) is for $t\geq 0^+$ a C^∞-trajectory which is contained in M_f^*, for all $\underline{x}_0\in\mathbb{R}^n$. Thus, we may state:

Remark (7.23): Let $(F,G)\in\mathbb{R}^{nxn}x\mathbb{R}^{nxn}$ be an e.r. pair. For ∀ initial vector $\underline{x}_0\in\mathbb{R}^n$ the general trajectory $\underline{x}(t,\underline{x}_0)$ is for $\forall\ t\geq 0^+$ a C^∞-trajectory which is contained in M_f^* space of (F,G). The space M_f^* has thus a strong attractivity property for all trajectories initiated by $\forall\ \underline{x}_0\in\mathbb{R}^n$. □

7.5.6. Dynamic characterisation of the generalised invariant subspaces of a general pair (F,G)

In the study of the geometry of subspaces of the domain of a general pair $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$, a number of important notions of invariance have been introduced i.e. the notions of (G,F)-i.s., (F,G)-i.s., c-(F,G)-i.s., $(\mathcal{W},\mathcal{U})$-p.i.s., e-(F,G)-i.s., e-(G,F)-i.s., c.e.-(G,F)-i.s. and right-(F,G)-annihilating spaces. These subspaces have been characterised in terms of the set of strict equivalence invariants of $(F,G)/\mathcal{V}$ and thus in terms of decompositions of the type (7.38a); such decompositions, together with the results of the previous section yield the following dynamic characterisation of the various types of generalised invariant subspaces.

Theorem (7.16): Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be a general pair, $\mathcal{V}\subset\mathbb{R}^n$ a subspace and $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ be the differential systems associated with (F,G). Then,

(i) If $\mathcal{V}$ is a (G,F)-i.s., then for $\mathcal{S}(F,G)$ is a C^∞-h.s., and for $\hat{\mathcal{S}}(F,G)$ a D_B'-h.s.

(ii) If $\mathcal{V}$ is a (F,G)-i.s., then for $\mathcal{S}(F,G)$ is a D_B'-h.s. and for $\hat{\mathcal{S}}(F,G)$ a C^∞-h.s.

(iii) If $\mathcal{V}$ is a c-(F,G)-i.s., then for both $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ it is a C^∞-h.s.

(iv) If $\mathcal{V}$ is a $(\mathcal{W},\mathcal{U})$-p.i.s., then for both $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ it is a D_B'-h.s. □

The proof of the above result follows from the decomposition properties of the subspaces involved, in terms of the fundamental subspaces of the type V_o, V_α, V_∞, V_ε, from the duality of the pencils $(F,G)/V$, $(G,F)/V$ and from the characterisation of the elementary subspaces V_o, V_α, V_∞, V_ε in terms of the solutions restricted to them. Similar arguments may be used to provide a dynamic characterisation of the maximal subspaces R^*, K^*, L^*, J^*, T^*, W^* defined before. We summarise as follows:

Theorem (7.17): Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$ be a general pair, $I(F,G)$ the set of strict equivalence invariance and let

$$\mathbb{R}^n = V_\zeta \oplus V_\varepsilon \oplus V_o \oplus V_\alpha \oplus V_\infty \tag{7.94}$$

be a decomposition of $\mathbb{R}^n$ implied by the set $I(F,G)$.

(i) The maximal right annihilating space $R^*=V_\varepsilon$ is the maximal C^∞-r.s. for both systems $S(F,G)$, $\hat{S}(F,G)$.

(ii) The maximal (G,F)-invatiant subspace $T^*=V_\varepsilon\oplus V_o\oplus V_\alpha$ is the maximal C^∞-h.s. for $S(F,G)$.

(iii) The maximal (F,G)-invariant subspace J^*, $J^*=V_\varepsilon\oplus V_\infty\oplus V_\alpha$ is the maximal C^∞-h.s. for $\hat{S}(F,G)$.

(iv) The maximal almost (F,G)-right annihilating space K^*, $K^*=V_\varepsilon\oplus V_\infty$, is the maximal D_B'-r.s. of $S(F,G)$.

(v) The maximal almost (G,F)-right annihilating space L^*, $L^*=V_\varepsilon\oplus V_o$, is the maximal D_B'-r.s. of $\hat{S}(F,G)$.

(vi) The maximal complete (F,G)-invariant subspace W^*, $W^*=V_\varepsilon\oplus V_\alpha$, is the maximal C^∞-h.s. for both systems $S(F,G)$, $\hat{S}(F,G)$.

(vii) The subspace $Q^*=T^*+J^*$ is the maximal (W,U)-p.i.s.; Q^* is also the maximal D_B'-h.s. for both systems $S(F,G)$, $\hat{S}(F,G)$ and may be expressed as $Q^*=V_\infty\oplus V_\varepsilon\oplus V_\alpha\oplus V_o$.

□

The proof is readily established from the previous results. We may now use Theorem (7.17) to characterise the properties of the spaces $\mathcal{C}$, $\mathcal{C}^*$ associated with a differential system $\mathcal{S}(F,G)$.

<u>Corollary (7.20)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$, $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ be the associated differential systems and let $\mathcal{C}$, $\hat{\mathcal{C}}$ and $\mathcal{C}^*$, $\hat{\mathcal{C}}^*$ be the initial spaces and redundancy spaces of $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ respectively. Then,

(i) $\mathcal{C}=\hat{\mathcal{C}}=\mathcal{Q}^*$

(ii) $\mathcal{C}^*=\hat{\mathcal{C}}^*=\mathbb{R}^n-\mathcal{Q}^*$

(iii) $\mathcal{T}^*\subset\mathcal{C}$ is the maximal space of C^∞-c.i.v. for $\mathcal{S}(F,G)$ and $\mathcal{J}^*\subset\hat{\mathcal{C}}^*$ is the maximal space of C^∞-c.i.v. for $\hat{\mathcal{S}}(F,G)$.

□

The properties of the $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ defined by the terms regular, singular C^∞-reachable, D_B'-reachable may now be characterised in terms of the set of invariants of (F,G), or equivalently in terms of maximal invariant subspaces as:

<u>Corollary (7.21)</u>: Let $(F,G)\in\mathbb{R}^{mxn}x\mathbb{R}^{mxn}$, $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ be the associated systems, $\mathcal{I}(F,G)=\{\mathcal{I}_r;\mathcal{I}_c;\mathcal{D}_o,\mathcal{D}_\alpha,\mathcal{D}_\infty\}$ the set of strict equivalence invariants of $sF-\hat{s}G$. Then,

(i) $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ are regular, iff either of the following equivalent conditions hold true.

(a) $\mathcal{I}_r=\mathcal{I}_r^o$ (a set of zero r.m.i., or $\emptyset$) and $\mathcal{I}_c=\emptyset$.

(b) $\mathcal{Q}^*=\mathbb{R}^n$ and $\mathcal{R}^*=\{0\}$.

(ii) $\mathcal{S}(F,G)$, $\hat{\mathcal{S}}(F,G)$ are C^∞-reachable, iff either of the following equivalent conditions hold true.

(a) $\mathcal{I}_r=\mathcal{I}_r^o$ and $\mathcal{D}_o=\emptyset$, $\mathcal{D}_\alpha=\emptyset$, $\mathcal{D}_\infty=\emptyset$.

(b) $\mathcal{R}^*=\mathbb{R}^n$

(iii) $\mathcal{S}(F,G)$ is D'_B-reachable, iff either of the following equivalent conditions hold true.

(a) $\mathcal{I}_r=\mathcal{I}^o_r$, $\mathcal{D}_o=\emptyset$, $\mathcal{D}_\alpha=\emptyset$

(b) $\mathcal{K}^*=\mathbb{R}^n$

(iv) $\hat{S}(F,G)$ is D'_B-reachable , iff either of the following equivalent conditions hold true.

(a) $\mathcal{I}_r=\mathcal{I}^o_r$, $\mathcal{D}_\alpha=\emptyset$, $\mathcal{D}_\infty=\emptyset$

(b) $\mathcal{L}^*=\mathbb{R}^n$

□

The expression i.e. $\mathcal{D}_\alpha=\emptyset$, in the above result means that the set $\mathcal{I}(F,G)$ has no $\mathcal{D}_\alpha$ subset in it.

7.6. Conclusions

The aim of this chapter was to provide a treatment of the subspaces of the domain of a general pair (F,G) from the geometric, algebraic and dynamic viewpoint. The basis for this work has been the characterisation of subspaces $\mathcal{V}$ in terms of the strict equivalence invariants of $(F,G)/\mathcal{V}$ restriction pencil. The motivation for this study has been the desire to detach the well developed geometric theory of regular state space systems [Won. -1],[Will. -1] e.t.c. from the specific content of regular state space systems and thus make it applicable to more general descriptions, such as the family of extended state space systems.

The notion of regular invariant realisations allows the translation of the results derived for $\mathcal{S}(F,G)$ systems, back to the framework of regular state space theory. In fact the notions of (F,G) is right annihilating spaces, are equivalent to the standard notions of (A,B)-invariant subspaces, controllability subspaces respectively.

The notion of $(\mathcal{W},\mathcal{U})$-p.i.s. corresponds to the standard notion of almost (A,B)-invariant subspace, while the almost (F,G)-right annihilating spaces correspond to the almost controllability subspace notion.

The treatment given here is by no means complete; although the "almost controlled invariance" and "almost reachability" properties with smooth trajectories may be transferred as a characterisation on the p.m.r.i.r. $(N,NA;\tilde{\mathcal{V}})$, which corresponds to a triple $(F,G;\mathcal{V})$, the trajectories do not belong to the domain of (F,G). It is an open question whether such important norions may be transferred on a general $\mathcal{S}(F,G)$ description. It seems that they are related to the specific property that (F,G) is entirely right singular.

The extension of the piecewise arithmetic progression sequences to the case of singular pair (F,G) provides the means for the computation of the Kronecker form by inspection. It is believed that the results of chapter (4) and (5) may provide a systematic procedure for the computation of $\mathcal{K}^*$, $\mathcal{L}^*$ along similar lines to those given for $\mathcal{R}^*$.

CHAPTER 8:

Topological and Relativistic Aspects of Matrix Pencils and S(F,G) Systems

CHAPTER 8: TOPOLOGICAL AND RELATIVISTIC ASPECTS OF MATRIX PENCILS AND $S(F,G)$ SYSTEMS

8.1 Introduction

In the previous chapters a detailed exposition of the algebraic, geometric and dynamic aspects of the subspaces of the domain of a pair (F,G), or of the associated differential system $S(F,G)$ has been given. An implicit assumption in this study has been that the pair (F,G) is fixed. However, uncertainty in the parameters of (F,G) is always inherent, whenever (F,G) emerges as the pair describing the differential equations of linear systems problem, or whenever it is the result of some previous stage of compuatations. Thus, the need for the study of perturbation properties of the pair (F,G) is of considerable importance and deserves special attention.

A vast amount of literature exists on the perturbation theory of the generalised eigenvalue-eigenvector problem (see for instance [Wilk. -1], [Ste.-1,2] etc.). Most of these results deal with the perturbation properties of the root range Φ of (F,G); this area of research is still open in the direction of obtaining stronger bounds and extending the existing results to singular pairs. The robustness problem of the complete set of invariants of (F,G), as well as the study of perturbation properties of the generalised invariant subspaces is also open. On the perturbation aspects of generalised invariant subspaces very few results are known with the noticable exception of the work of Stewart[Ste. -1,2] and Van Dooren [Van Do. -3] on deflating subspaces. The study of the generic properties of the pair (F,G) is also open; the results in this area are few with the exception on those dealing with the generic values of the root range Φ [Hir. & Sm. -1] and those given in [Won. -1] on the generic values of c.m.i., r.m.i. on entirely right, left singular pairs.

The study of properties of a general pair (F',G') in a ball centered at some nominal pair (F,G), is the subject of the perturbation theory. A similar in the formulation, but of entirely different nature problem, is the study of properties (that stem from the set of strict equivalence invariants) of a regular pair $(\tilde{F},\tilde{G})$ in the bilinear orbit $\mathcal{E}_B(F,G)$ of a given pair (F,G). Given that the strict equivalence invariants are "space invariants" (invariant under coordinate transformations in the domain and codomain of (F,G) and bilinear transformations express transformations in the frequency domain (coordinate transformations on the Riemann sphere), this study may be termed as "space-frequency relativistic". This study is of theoretical importance, since it provides a classification of the geometric and dynamic aspects of (F,G) to those which are "space-frequency invariant", and those which are "relativistic" i.e. deppend on the frequency coordinate frame. Apart from its theoretical importance, from the conceptual viewpoint, this study provides the means for the constraction of "convenient dual problems" in linear systems, and as it has been discussed in Chapter (6) may also yield "convenient descriptions" from the computational viewpoint (see problem of assignment of condition number).

The chapter is divided into two parts. The first deals with the topological aspects and the second with the "relativistic". The topological results are of a preliminary nature and aims at two directions: first to provide a unifying framework for the study of perturbation aspects by introducing different metric topologies on the pairs (F,G) and second to establish the links between the new metrics and the already known results. The angle metric is connected with the work of Sun Ji-guang [Sun J.g. -1], and the deflating subspaces of Stewart are shown to be equivalent to the notion of e.d. subspace examined previously. The second part of the chapter deals with the classification of the geometric notions of invariant subspaces of

(F,G), as well as the classification of their dynamic counterparts. The results in this latter part are based on the theory of E_{H-B} equivalence developed in Chapter (6).

8.2 Classification of pairs, Characteristic spaces

We start off by giving some useful definitions and notation. We shall denote by $L_{m,n} = \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$.

Definition (8.1): For a given pair $L = (F,G) \in L_{m,n}$ we define:

(i) A flat matrix representation of L: $[L]_f = [F \,\vdots\, -G] \in \mathbb{R}^{m\times 2n}$ (8.1)

(ii) A sharp matrix representation of L: $[L]_s = \left[\begin{array}{c} F \\ \hline -G \end{array}\right] \in \mathbb{R}^{2m\times n}$ (8.2)

□

Definition (8.2): Let $L = (F,G) \in L_{m,n}$, then L will be called nondegenerate

$$\begin{cases} \text{rank } [L]_f = m, \text{ if } m \leq n \\ \text{rank } [L]_s = n, \text{ if } m \geq n \end{cases} \quad (8.3)$$

□

Remark (8.1): Degeneracy, in the case of flat description implied zero row minimal indices, whereas in the case of sharp description implies zero column minimal indices.

Remark (8.2): If m = n, then nondegeneracy implies that L is a regular pair. By using these definitions we can express the notion of strict equivalence as follows:

Definition (8.3): Let $L, L' \in L_{m,n}$

(a) L,L' will be called strict equivalent (LE_HL'), iff $\exists Q \in \mathbb{R}^{n\times n}, R \in \mathbb{R}^{m\times m}$, $|P|, |Q| \neq 0$, such that:

$$\text{(i)} \qquad [L']_f = R[L]_f \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} \qquad (8.4)$$

or equivalently

$$\text{(ii)} \qquad [L']_s = \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix} [L]_s Q \tag{8.5}$$

(b) L,L' will be called <u>left strict equivalent</u>, $L\mathcal{E}_{\mathcal{H}}^{\ell}L'$, iff

$$[L']_f = R[L]_f, \quad R \in \mathbb{R}^{m\times m}, \quad |R| \neq 0 \tag{8.6}$$

and shall be called <u>right strict equivalent</u>, $L\mathcal{E}_{\mathcal{H}}^{r}L'$, iff

$$[L']_s = [L]_s Q, \quad Q \in \mathbb{R}^{n\times n}, \quad |Q| \neq 0 \tag{8.7}$$

□

For the pair $L = (F,G) \in \mathcal{L}_{m,n}$ we can define four vector spaces as follows:

$$\begin{aligned} \mathcal{X}_f^r(L) &\triangleq \text{rowspan}_{\mathbb{R}}[L]_f \quad , \quad \mathcal{X}_f^c(L) \triangleq \text{colspan}_{\mathbb{R}}[L]_f \quad , \\ \mathcal{X}_s^r(L) &\triangleq \text{rowspan}_{\mathbb{R}}[L]_s \quad , \quad \mathcal{X}_s^c(L) \triangleq \text{colspan}_{\mathbb{R}}[L]_s \end{aligned} \tag{8.8}$$

Then it is clear, that:

$\forall\ L,L' \in \mathcal{L}_{m,n}$: $L\mathcal{E}_{\mathcal{H}}^{\ell}L' \quad \mathcal{X}_f^r(L) = \mathcal{X}_f^r(L')$ and $\mathcal{X}_s^r(L) = \mathcal{X}_s^r(L')$

which means that $\mathcal{X}_f^r(L)$, $\mathcal{X}_s^r(L)$ are invariant for $\mathcal{E}_{\mathcal{H}}^{\ell}$-equivalence, but $\mathcal{X}_f^c(L)$, $\mathcal{X}_s^c(L)$ are not always invariant for $\mathcal{E}_{\mathcal{H}}^{\ell}$-equivalence.

Also, $\forall\ L,L' \in \mathcal{L}_{m,n}$: $L\mathcal{E}_{\mathcal{H}}^{r}L' \quad \mathcal{X}_f^c(L) = \mathcal{X}_f^c(L')$ and $\mathcal{X}_s^c(L) = \mathcal{X}_s^c(L')$ which also means that $\mathcal{X}_f^c(L)$, $\mathcal{X}_s^c(L)$ are invariant under $\mathcal{E}_{\mathcal{H}}^{r}$-equivalence, but $\mathcal{X}_f^r(L)$, $\mathcal{X}_s^r(L)$ are not always invariant under $\mathcal{E}_{\mathcal{H}}^{r}$-equivalence. □

Next we shall give a classification of the pairs, according to their dimensions, and we shall examine the invariants properties of the four vector spaces which are defined on each pair of the corresponding class.

Let $L = (F,G) \in \mathcal{L}_{m,n}$. The following classification is a simple consequence of the definitions and observations, so far.

Table (8.1) demonstrates the type of subspace, with invariant properties, we have to chose according to the dimensionality (relationship between m and n) of the particular family of pairs L.

Table (8.1)

$L = (F\text{-}G)$
$\in$
$L_{m,n}$

Dimensions	$X^{r}_{f}(L)$	$X^{r}_{s}(L)$	$X^{c}_{f}(L)$	$X^{c}_{s}(L)$	Invariant spaces
$m \leq n/2$	E^{ℓ}_{H}-invariant	E^{ℓ}_{H}-invariant	$X^{c}_{f}(L) \equiv \mathbb{R}^{2m}$ E_{H}-invariant	$X^{c}_{s}(L) \equiv \mathbb{R}^{2m}$ E_{H}-invariant	$X^{r}_{f}(L), X^{r}_{s}(L)$
$2n \geq m > n/2$	E^{ℓ}_{H}-invariant	$X^{r}_{s}(L) \equiv \mathbb{R}^{2m}$ E_{H}-invariant	$X^{c}_{f}(L) \equiv \mathbb{R}^{2m}$ E_{H}-invariant	E^{r}_{H}-invariant	$X^{r}_{f}(L), X^{c}_{s}(L)$
$n \leq m/2 < 2m$	$X^{r}_{f}(L) \equiv \mathbb{R}^{2n}$ E_{H}-invariant	$X^{r}_{s}(L) \equiv \mathbb{R}^{n}$ E_{H}-invariant	E^{r}_{H}-invariant	E^{r}_{H}-invariant	$X^{c}_{f}(L), X^{c}_{s}(L)$
$2m \geq n > m/2$	E^{ℓ}_{H}-invariant	$X^{r}_{s}(L) \equiv \mathbb{R}^{n}$ E_{H}-invariant	$X^{c}_{f}(L) \equiv \mathbb{R}^{m}$ E_{H}-invariant	E^{r}_{H}-invariant	$X^{r}_{f}(L), X^{c}_{s}(L)$
Regular, $m=n$	E^{ℓ}_{H}-invariant	$X^{r}_{s}(L) \equiv \mathbb{R}^{n}$ E_{H}-invariant	$X^{c}_{f}(L) \equiv \mathbb{R}^{n}$ E_{H}-invariant	E^{r}_{H}-invariant	$X^{r}_{f}(L), X^{c}_{s}(L)$

Definition (8.4): For a non-degenerate pair $L = (F, G) \in \mathcal{L}_{m,n}$ we define the normal-characteristic space, $\mathcal{X}_N(L)$, as follows:

(i) $\mathcal{X}_N(L) \triangleq \mathcal{X}_f^r(L)$, if $m \leq n$, which is $\mathcal{E}_H^\ell$-invariant

(ii) $\mathcal{X}_N(L) \triangleq \mathcal{X}_s^c(L)$, if $m > n$, which is $\mathcal{E}_H^r$-invariant □

A straight consequence of (i) is that for regular pairs $L = (F,G) \in \mathcal{L}_{n,n}$, we have that $\mathcal{X}_N(L) = \mathcal{X}_f^r(L)$. Furthermore, note that:

$$\mathcal{X}_f^r(L) = \text{rowspan}[L]_f = \text{colspan}[L]_f^t, \text{ and } \mathcal{X}_s^c(L) = \text{colspan}[L]_s = \text{rowspan}[L]_s^t, \quad (8.9)$$

Let $L = (F, G) \in \mathcal{L}_{m,n}$, $m \leq n$ non degenerate pair, then $\text{rank}[L]_f = m$ and $\mathcal{X}_N(L) = \mathcal{X}_f^r(L) = \text{rowspan}[L]_f$ and thus $\dim \mathcal{X}_N(L) = m$. Therefore, we can identify $\mathcal{X}_N(L)$ with a point of the Grassmann Manifold $G(m, \mathbb{R}^{2n})$. Note that if $L_1 \mathcal{E}_H^\ell L$, then $\mathcal{X}_N(L_1)$ represents the same point on $G(m, \mathbb{R}^{2n})$ as $\mathcal{X}_N(L)$. □

Let $\mathcal{L}_{n,n}^r$ be the set of regular pairs. $\forall$ $L = (F, G) \in \mathcal{L}_{n,n}^r$ the equivalent class $\mathcal{E}_H^\ell(L) = \{L_1 \in \mathcal{L}_{n,n}^r : [L_1]_f = R[L]_f, \ R \in \mathbb{R}^{n\times n}, \ |R| \neq 0\}$ can be considered as representing a point on $G(n, \mathbb{R}^{2n})$; this point may be identified with a $\mathcal{X}_N(L_1)$, $L_1 \in \mathcal{E}_H^\ell(L)$. Note that each point of $G(m, \mathbb{R}^{2n})$ does not represent a pair, but a left-equivalent calss of pairs. We shall denote by Σ_1^L the set of all distinct generalised eigenvalues (finite and infinite) of the pair $L \in \mathcal{L}_{n,n}^r$ and by Σ_2^L the set of normalized generalised eigenvectors (Jordan vectors included) defined $\forall$ $\lambda \in \Sigma_1^L$.

Let $\mathcal{Q} = \{(\Sigma_1^L, \Sigma_2^L) : \forall \ L \in \mathcal{L}_{n,n}^r\}$, we define the function,

$$f: \mathcal{L}_{n,n}^r \longrightarrow \mathcal{Q}: L \to f(L) \triangleq (\Sigma_1^L, \Sigma_2^L) \in \mathcal{Q} \quad (8.10)$$

It is obvious that $f^{-1}(\{(\Sigma_1^L, \Sigma_2^L)\}) = \mathcal{E}_H^\ell(L)$, which means that f is $\mathcal{E}_H^\ell$-invariant, thus from the eigenvalue-eigenvector point of view, it is reasonable to identify the hole class $\mathcal{E}_H^\ell(L)$ as a point of the Grassmann manifold. The study of Topological properties of ordered regular pairs is intimaterly related with the generalised eigenvalue problem. The analysis above demonstrates that such properties may be studied on the appropriate Grasmann manifold. We may introduce a metric topology on the manifold in different ways. One such new metric is discussed next.

8.3 The angle metric

In the following we shall consider Grassmann manifolds which may be associated with elements of $\mathcal{L}_{m,n}$. In the case where $m \leq n$, the manifold $G(m, \mathbb{R}^{2n})$ is considered, whereas in the case of $m \geq n$ the manifold $G(n, \mathbb{R}^{2m})$ is the subject of examination. The metric defined in this section on the above manifolds, clearly covers the regular case. The relationship of the two Grassmann manifolds with pairs of $\mathcal{L}_{m,n}$ is clarified by the following remark.

Remark (8.3): (i) Let $\mathcal{V} \in G(m, \mathbb{R}^{2n})$. There exists an $\mathcal{E}_H^{\ell}$-equivalence class from $\mathcal{L}_{m,n}$ $(m \leq n)$, represented by an $L = (F, G) \in \mathcal{L}_{m,n}$ nondegenerate, such that $\mathcal{V} = \text{rowspan}[L]_f$. (ii) Let $\mathcal{V} \in G(n, \mathbb{R}^{2m})$. There exists an $\mathcal{E}_H^{r}$-equivalence class from $\mathcal{L}_{m,n}$ $(m \geq n)$, represented by an $L = (F, G) \in \mathcal{L}_{m,n}$ nondegenerate, such that $\mathcal{V} = \text{colspan}[L]_s$. □

Definition (8.5): A real valued function: $d(\cdot,\cdot): G^2(m, \mathbb{R}^{2n}) \to \mathbb{R}_o^+$, $(m \leq n)$ is called an orthogonal invariant metric (O.I.M.). If $\forall\ \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \in G(m, \mathbb{R}^{2n})$ (we may assume that $\mathcal{V}_1 = \mathcal{X}_N(L_1)$, $\mathcal{V}_2 = \mathcal{X}_N(L_2)$, $\mathcal{V}_3 = \mathcal{X}_N(L_3)$, $L_i \in \mathcal{L}_{m,n}$, $i = 1,2,3$ and non degenerate) it satisfies the following properties:

(i) $d(\mathcal{X}_N(L_1), \mathcal{X}_N(L_2)) \geq 0$, $d(\mathcal{X}_N(L_1), \mathcal{X}_N(L_2)) = 0$ iff $L_1 \mathcal{E}_H^{\ell} L_2$

(ii) $d(\mathcal{X}_N(L_1), \mathcal{X}_N(L_2)) = d(\mathcal{X}_N(L_2), \mathcal{X}_N(L_1))$

(iii) $d(\mathcal{X}_N(L_1), \mathcal{X}_N(L_2)) \leq d(\mathcal{X}_N(L_1), \mathcal{X}_N(L_3)) + d(\mathcal{X}_N(L_3), \mathcal{X}_N(L_2))$

(iv) $d\left(\text{colspan}\left[[L_1]_f^t U\right], \text{colspan}\left[[L_2]_f^t U\right]\right) = d(\text{colspan}[L_1]_f^t, \text{colspan}[L_2]_f^t)$

where $U \in \mathbb{R}^{m \times m}$ is an orthogonal matrix. □

Similarly we can define an orthogonal invariant metric on $G(n, \mathbb{R}^{2m})$. First we summarise some already known metrics for $G(m, \mathbb{R}^{2n})$, $m \leq n$, and then we introduce a new metric, the angle metric. Finally the relationship between the new metric and the previously defined ones are examined.

<u>Theorem (8.1)</u>: [Sun J.G. -1]. For any two points $\mathcal{V}_1, \mathcal{V}_2 \in G(m, \mathbb{R}^{2n})$, such that $\mathcal{V}_1 = \mathcal{X}_f^{r}(L_1) = \text{colspan}[L_1]_f^t$, $\mathcal{V}_2 = \mathcal{X}_f^{r}(L_2) = \text{colspan}[L_2]_f^t$ we define $X_1 = ([L_1]_f^t[L_1]_f)^{-1/2} \cdot [L_1]_f^t$, $X_2 = ([L_2]_f^t[L_2]_f)^{-1/2} \cdot [L_2]_f^t$, then

$$\ell(\mathcal{V}_1, \mathcal{V}_2) = \arccos[\det X_1 X_2^t X_2 X_1^t]^{1/2} \tag{8.11}$$

and

$$d_\ell(\mathcal{V}_1, \mathcal{V}_2) = \sin \ell(\mathcal{V}_1, \mathcal{V}_2) \tag{8.12}$$

are O.I.M. on $G(m, \mathbb{R}^{2n})$ □

These metrics have been used by Ji-G-Sun on his work on perturbation analysis for the generalised eigenvalue-eigenvector problem. The above metrics have also been related [Sun. J.g. -1,2] to the other metrics on the Grassmann manifold (e.g. "gap" between subspaces, $\sin\theta(X,Y)$). Let's denote by $\underset{\rightarrow}{\mathcal{L}}_{m,n} = \{L = (F,G) \in \mathcal{L}_{m,n},\ m \leq n,\ L \text{ nondegenerate}\}$ and by $\underset{\rightarrow}{\mathcal{L}}_{m,n}/\mathcal{E}_{\mathcal{H}}^{\ell}$ the quotient $\mathcal{E}_{\mathcal{H}}^{\ell}$-orbit we have seen that there is an one to one map φ defined by:

$$\varphi: \underset{\rightarrow}{\mathcal{L}}_{m,n}/\mathcal{E}_{\mathcal{H}}^{\ell} \longrightarrow G(m, \mathbb{R}^{2n}) : \mathcal{E}_{\mathcal{H}}^{\ell}(L) \longrightarrow \varphi(\mathcal{E}_{\mathcal{H}}^{\ell}(L)) \overset{\Delta}{=} \mathcal{X}_N(L) = \mathcal{X}_f^{r}(L) \overset{\Delta}{=} \mathcal{V}_L \in G(m, \mathbb{R}^{2n}) \tag{8.13}$$

withy every m-dimensional subspace $\mathcal{V}_L$ of $\mathbb{R}^{2n}$ we may always associate an one-dimensional subspace of the vector space $\Lambda^m(\mathbb{R}^{2n})$ [Mar. -1], or in other words a point on the Grassmann variety $\Omega(m,2n)$ of the projective space $\mathbb{P}^v(\mathbb{R})$, $v = \binom{2n}{m} - 1$.

These one-dimensional subspaces of $\Lambda^m(\mathbb{R}^{2n})$ consist of decomposable multivectors, they are called Grassmann representatives of the corresponding subspace $\mathcal{V}_L$ and they are denoted by $\underline{g}(\mathcal{V}_L)$ It is known that if $\underline{x}, \underline{x}' \in \Lambda^m(\mathbb{R}^{2n})$ are two Grassmann representatives of $\mathcal{V}_L$ then, $\underline{x}' = \lambda \underline{x}$, where $\lambda \in \mathbb{R} - \{0\}$. □

<u>Definition (8.6)</u>: Let $L = (F, G) \in \underline{\mathcal{L}}_{m,n}$ and $\mathcal{V}_L \in G(m, \mathbb{R}^{2n})$ be the linear subspace of $\mathbb{R}^{2n}$ which corresponds to the equivalent class $\mathcal{E}_{\mathcal{H}}^{\ell}(L)$. If $\underline{x} \in \Lambda^m(\mathbb{R}^{2n})$ is a Grassmann representative of $\mathcal{V}_L$ and $\underline{x} = [x_o, x_1, \dots x_v]^t$, $v = \binom{2n}{m} - 1$, then we define as the <u>normal Grassmann representative</u> of $\mathcal{X}_L$ the decomposable multivector $\underline{\tilde{x}}$ which is given by:

$$\underline{\tilde{x}} \overset{\Delta}{=} \begin{cases} \dfrac{\text{sign}(x_v)}{\|\underline{x}\|_2} \cdot \underline{x}, & \text{if } x_v \neq 0 \\[2ex] \dfrac{\text{sign}(x_i)}{\|\underline{x}\|_2} \cdot \underline{x}, & \text{if } x_v = 0,\ x_i \text{ is the first non zero component of } \underline{x} \end{cases} \tag{8.14}$$

where $\text{sign}(x_i) = \begin{cases} 1, & \text{if } x_i > 0 \\ -1, & \text{if } x_i < 0 \end{cases}$

□

<u>Proposition (8.1)</u>: The normal Grassmann representative $\underline{\tilde{x}}$ of a subspace $\mathcal{V}_L \in G(m, \mathbb{R}^{2n})$ is unique and $\|\underline{\tilde{x}}\|_2 = 1$. ($\|\cdot\|_2$ denotes the Euclidean norm).

<u>Proof</u>

Let $\underline{x}' \in \Lambda^m(\mathbb{R}^{2n})$ be another Grassmann representative of $\mathcal{V}_L$. Then clearly $\underline{x}' = \lambda\underline{x}$, $\lambda \in \mathbb{R} - \{0\}$ and $\underline{\tilde{x}}'$ (the normal Gr. representative of $\underline{x}'$) is given by:

$$\underline{\tilde{x}}' = \begin{cases} \dfrac{\text{sign}(x'_v)}{\|\underline{x}'\|_2} \cdot \underline{x}', & \text{if } x_v = \lambda x_v \neq 0 \\[2ex] \dfrac{\text{sign}(x'_i)}{\|\underline{x}'\|_2} \cdot \underline{x}', & \text{if } x_v = \lambda x_v = 0 \text{ and } x'_i = \lambda x_i \text{ is the first non zero component of } \underline{x}' \end{cases}$$

$$= \begin{cases} \dfrac{\lambda \text{sign}\lambda \text{sign}(x_v)}{|\lambda| \cdot \|\underline{x}\|_2} \cdot \underline{x}, & \text{if } x_v \neq 0 \\[2ex] \dfrac{\lambda \text{sign}\lambda \text{sign}(x_i)}{|\lambda|\ \|\underline{x}\|_2} \cdot \underline{x}, & \text{if } x_v = 0 \text{ and } x_i \text{ the first non zero component of } \underline{x}. \end{cases} = \underline{\tilde{x}}$$

(because $\dfrac{\lambda\text{sign}(\lambda)}{|\lambda|} = 1$).

So $\tilde{\underline{x}}$ is a unique representative of the class of all Grassmann representatives of the subspace $\mathcal{V}_L$. The property $\|\tilde{\underline{x}}\|_2 = 1$ is obvious. □

Definition (8.7): Let $\mathcal{V}_{L_1}, \mathcal{V}_{L_2} \in G(m, \mathbb{R}^{2n})$, then we can define an angle $\sphericalangle(\mathcal{V}_{L_1}, \mathcal{V}_{L_2})$ between them as follows:

$$\sphericalangle(\mathcal{V}_{L_1}, \mathcal{V}_{L_2}) \triangleq \arccos |<\tilde{\underline{x}}_1, \tilde{\underline{x}}_2>| \tag{8.15}$$

where $\tilde{\underline{x}}_1$, $\tilde{\underline{x}}_2$ are the normal Grassmann representatives of $\mathcal{V}_{L_1}$, $\mathcal{V}_{L_2}$ respectively and $<\cdot,\cdot>$ denotes inner product. □

Proposition (8.2): The angle defined above is an orthogonal invariant metric on the Grassmann manifold $G(m, \mathbb{R}^{2n})$.

Proof

Properties (i), (ii) follow readily, property (iii) follows from the corresponding inequality for angles between vectors, which is a consequence of the Δ-inequality on the unit sphere.

To prove (iv), let $\mathcal{V}_{L_i} = \text{colspan}[L_i]_f^t$, $i = 1,2$ and U an mxm non singular matrix with $|U| = \lambda$. Then $C_m([L_i]_f^t U) = C_m([L_i]_f^t)|U| = \lambda \cdot C_m([L_i]_f^t)$, $i = 1,2$, which means that $\sphericalangle([L_1]_f^t U, [L_2]_f^t U) = \sphericalangle([L_1]_f^t, [L_2]_f^t)$, if we suppose that U is orthogonal then clearly it is an orthogonal invariant metric.

This metric angle on $G(m, \mathbb{R}^{2n})$ is related to the standard metrics (8.11), (8.12) as it is shown below.

Proposition (8.3): $\forall\ \mathcal{V}_{L_1}, \mathcal{V}_{L_2} \in G(m, \mathbb{R}^{2n})$ we have

$$\begin{aligned} &(a) \quad \sphericalangle(\mathcal{V}_{L_1}, \mathcal{V}_{L_2}) = \ell\,(\mathcal{V}_{L_1}, \mathcal{V}_{L_2}) \\ &(b) \quad \sin\{\sphericalangle(\mathcal{V}_{L_1}, \mathcal{V}_{L_2})\} = d_\ell(\mathcal{V}_{L_1}, \mathcal{V}_{L_2}) \end{aligned} \tag{8.16}$$

Proof

By using the Binet-Cauchy theorem [Mar. & Min. -1], we have that

$$\det[X_1X_2^tX_2X_1^t] = C_m[X_1X_2^tX_2X_1^t] = C_m(X_1)\, C_m(X_2^t)\, C_m(X_2)\, C_m(X_1^t) =$$

$= C_m(X_1)\cdot C_m^t(X_2)\, C_m(X_2)\cdot C_m^t(X_1)$. From theorem (8.1) we also have,

$$C_m(X_i') = [C_m^t([L_i]_f)C_m([L_i]_f)]^{-1/2}\cdot C_m^t([L_i]_f) = \frac{\underline{g}(V_{L_i})}{\|\underline{g}(V_{L_i})\|_2} = \varepsilon\, \underline{\tilde{x}}_i^t,\ \varepsilon = \pm 1,\ i = 1,2$$

Thus,

$\det[X_1X_2^tX_2X_1^t] = \underline{\tilde{x}}_1^t\underline{\tilde{x}}_2\underline{\tilde{x}}_2^t\underline{\tilde{x}}_1 = \{<\underline{\tilde{x}}_1,\underline{\tilde{x}}_2>\}^2$, given that

$(\det[X_1X_2^tX_2X_1^t])^{1/2} = |<\underline{\tilde{x}}_1,\underline{\tilde{x}}_2>|$, then

(i) $\arccos\, (\det[X_1X_2^tX_2X_1^t])^{1/2} = \arccos\, |<\underline{\tilde{x}}_1,\underline{\tilde{x}}_2>|$, or

$$\ell(V_{L_1},V_{L_2}) = \sphericalangle(V_{L_1},V_{L_2})$$

(ii) $\sin\{\sphericalangle(V_{L_1},V_{L_2})\} = \sin\ell(V_{L_1},V_{L_2}) = d_\ell(V_{L_1},V_{L_2})$ □

Remark (8.4): When $m = n$ this angle metric is suitable for the perturbation analysis for the generalised eigenvalue-eigenvector problem. □

8.4 The chordal distance and perturbation results

In the study of sensitivity of eigenvalues a standard metric (expressing distance between eigenvalues (finite and infinite)) is used, that is the chordal-distance, between points of the Riemann sphere. Because of its importance in the perturbation theory, a proper treatment of this distance is given first and some new properties are established. Using the chordal distance and some ideas from the work of Stewart [Ste. -1] a new perturbation result for a singe eigenvalue of a regular pair is derived.

The formula expressing the coordinates of the point A(z) of the Riemann sphere, in terms of the coordinates of the point z of the plane is examined first.

Proposition (8.4): The point $z = x+jy$ under stereographic projection corresponds to the point $A(z)$, of the sphere $\xi^2+\eta^2+(\zeta-1/2)^2 = 1/4$, with coordinates, $\xi = 1/2\cdot{}^{x|z|}/_{1+|z|^2}$, $\eta = 1/2\cdot{}^{y|z|}/_{1+|z|^2}$, $\zeta = {}^{|z|^2}/_{1+|z|^2}$

Proof

Since the projection of the point A lies on the line OZ then, $\xi = \lambda x$, $\eta = \lambda y$, where λ is some real constant. We shall find ζ in terms of $|z|$. Let's consider the cross section of the sphere cut by the plane passing through the points, O, P and Z (Fig. 8.1).

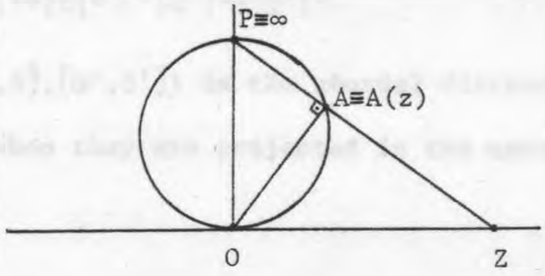

Figure (8.1)

The right triangles OPZ and OAZ are similar, the altitude of the triangle OAZ is equal to ζ and its hypotenuse is $|z|$. The segment OA is a leg of the triangle OAZ and the altitute of the triangle OPZ.

From the similarity of the triangles OPZ and OAZ we have

$$\zeta/\overline{OA} = \overline{OA}/\overline{OP}, \quad \zeta/\overline{OZ} = \overline{OA}/\overline{PZ}, \quad \overline{OZ} = |z|, \quad \overline{OP} = 1, \quad \overline{PZ} = \sqrt{1+|z|^2} \tag{8.17}$$

from which we find $\zeta = |z|^2/1+|z|^2$.

With the aid of the equation of the sphere it is easy to determine the value of $\lambda = 1/2\cdot|z|/1+|z|^2$ and after that that ξ and η.

□

Remark (8.5): We denote the distance between the points A(s) and A(s') by K(s,s'). With the aid of the formula of the previous proposition we can easily show that:

$$K(s,s') = |s-s'| \Big/ \sqrt{1+|s|^2} \cdot \sqrt{1+|s'|^2}, \quad K(s,\infty) = 1 \Big/ \sqrt{1+|s|^2} \qquad (8.18)$$

The quantity $K(s,s')$ is called the <u>chordal distance</u> between the points s and s'. In order to define a distance, which is suitable without making any distinction between finite and infinite eigenvalues, we shall identify the eigenvalue $s = \alpha/\beta$ with the point in the projective complex line defined by:

$$[\alpha,\beta] \overset{\Delta}{=} \{(\alpha,\beta) \neq (0,0) \ : \ \alpha/\beta = s\} \qquad (8.19)$$

Also we define on the projective complex line the metric K defined by:

$$\forall \ s=\alpha/\beta, \ s'=\alpha'/\beta' \Rightarrow K([\alpha,\beta],[\alpha',\beta']) = K(\alpha/\beta, \alpha'/\beta') =$$

$$= |\alpha\beta'-\alpha'\beta| \Big/ \sqrt{|\alpha|^2+|\beta|^2} \cdot \sqrt{|\alpha'|^2+|\beta'|^2} \qquad (8.20)$$

So the number $K([\alpha,\beta],[\alpha',\beta'])$ is the chordal distance between the points $s = \alpha/\beta$, $s' = \alpha'/\beta'$ when they are projected in the usual way onto the Riemann sphere. $\square$

<u>Proposition (8.5)</u>: For any two points $s,s' \in \mathbb{C} \cup \{\infty\}$ we have that $K(s,s') = K(1/s, 1/s')$.

<u>Proof</u>

Let $s = \alpha/\beta$, $s' = \alpha'/\beta'$ then:

$$K(s,s') = K([\alpha,\beta],[\alpha',\beta']) = |\alpha\beta'-\alpha'\beta| \Big/ \sqrt{|\alpha|^2+|\beta|^2} \cdot \sqrt{|\alpha'|^2+|\beta'|^2} \qquad (8.21)$$

$$K(1/s,1/s') = K([\beta,\alpha],[\beta',\alpha']) = |\beta\alpha'-\alpha\beta'| \Big/ \sqrt{|\beta|^2+|\alpha|^2} \cdot \sqrt{|\beta'|^2+|\alpha'|^2} \qquad (8.22)$$

So from (8.21), (8.22) we have $K(s,s') = K(1/s, 1/s')$ $\square$

<u>Proposition (8.6)</u>: K is invariant under orthogonal bilinear transformation.

<u>Proof</u>

We have to prove that, if

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}, \quad \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} = \begin{bmatrix} c & -d \\ c & c \end{bmatrix} \begin{bmatrix} \hat{\alpha}' \\ \hat{\beta}' \end{bmatrix}, \ c,d \in \mathbb{R}, \ c^2+d^2=1 \qquad (8.23)$$

then $K([\alpha,\beta],[\alpha',\beta']) = K([\hat{\alpha},\hat{\beta}],[\hat{\alpha}',\hat{\beta}'])$.

From (8.23) we have: $\alpha = c\hat{\alpha}-d\hat{\beta}$, $\beta = d\hat{\alpha}+c\hat{\beta}$, $\alpha' = c\hat{\alpha}'-d\hat{\beta}'$, $\beta' = d\hat{\alpha}'+c\hat{\beta}'$, so

$$K([\alpha,\beta],[\alpha',\beta'])=$$

$$=\left|(c\hat{\alpha}-d\hat{\beta})(d\hat{\alpha}'+c\hat{\beta}')-(c\hat{\alpha}'-d\hat{\beta}')(d\hat{\alpha}+c\hat{\beta})\right|/\sqrt{|c\hat{\alpha}-d\hat{\beta}|^2+|d\hat{\alpha}+c\hat{\beta}|^2}\cdot\sqrt{|c\hat{\alpha}'-d\hat{\beta}'|^2+|d\hat{\alpha}'+c\hat{\beta}'|^2} \quad (8.$$

The numerator of (8.24) after some simplifications can be written as

$$c^2(\hat{\alpha}\hat{\beta}'-\hat{\alpha}'\hat{\beta})+d^2(\hat{\alpha}\hat{\beta}'-\hat{\alpha}'\hat{\beta}) = \hat{\alpha}\hat{\beta}'-\hat{\alpha}'\hat{\beta} \tag{8.25}$$

Also, it is known that $\forall\ Z_1,Z_2 \in \mathbb{C} \Rightarrow |Z_1 \pm Z_2|^2 = |Z_1|^2+|Z_2|^2 \pm 2\mathrm{Re}(Z_1Z_2)$ (8.26) by using (8.26) the first term of the denominator of (8.24) can be written as
$|c\hat{\alpha}-d\hat{\beta}|^2+|d\hat{\alpha}+c\hat{\beta}|^2 = c^2|\hat{\alpha}|^2+d^2|\hat{\beta}|^2-2\mathrm{Re}(cd\hat{\alpha}\hat{\beta})+d^2|\hat{\alpha}|^2+c^2|\hat{\beta}|^2+2\mathrm{Re}(dc\hat{\alpha}\hat{\beta}) = |\hat{\alpha}|^2+|\hat{\beta}|^2$.

Also $|c\hat{\alpha}'-d\hat{\beta}'|^2+|d\hat{\alpha}'+c\hat{\beta}'|^2 = |\hat{\alpha}'|^2+|\hat{\beta}'|^2$, thus
$K([\alpha,\beta],[\alpha',\beta']) = K([\hat{\alpha},\hat{\beta}],[\hat{\alpha}',\hat{\beta}'])$. □

Next we derive a result for the sensitivity of a single eigenvalue of a regular pair (F,G). This result is based on the work by Stewart [Ste. -1] and provides a strong perturbation bound for the eigenvalues.

Let $F,G \in \mathbb{R}^{n\times n}$ such that the pencil $sF-G$ is regular. Let s_o be a single eigenvalue (multiplicity one) with $\underline{x}_o$ and $\underline{y}_o^t$ normalised ($\|\underline{x}_o\|_2 = \|\underline{y}_o^t\|_2 = 1$) right and left eigenvectors correspondingly, then $(s_oF-G)\underline{x}_o = 0$ and $\underline{y}_o^t(s_oF-G) = 0$. Also $s_o(\underline{y}_o^tF\underline{x}_o) = \underline{y}_o^tG\underline{x}_o$ and if we put $\underline{y}_o^tG\underline{x}_o = \alpha$, $\underline{y}_o^tF\underline{x}_o = \beta$ the point $s_o = \alpha/\beta$ can identified with the point $[\alpha,\beta]$ of the projective line.

<u>Proposition (8.7) [Ste. -1]</u>: Let the regular pair $(F,G) \in L^{r}_{n,n}$. Let $s_o = \alpha/\beta$ be a single eigenvalue for the pair (F,G) with $\underline{x}_o$, $\underline{y}_o^t$ normalised right and left eigenvectors correspondingly. For sufficiently small $G',F' \in \mathbb{R}^{n\times n}$ there is an eigenvalue s_o' for the pair $(F+F',-(G+G'))$ that can be identified with the point $[\alpha+\underline{y}_o^tF'\underline{x}_o,\ \beta+\underline{y}_o^tG'\underline{x}_o]$ of the projective line, except for terms of order $\|F'\|_2^2$, $\|G'\|_2^2$. □

Remark (8.6): According to this result, the sensitivity of s_o to perturbations in G and F will be measured in terms of the chordal distance by:

$$d(s_o,s_o') \cong K([\alpha,\beta],[\alpha+\underline{y}_o^t F'\underline{x}_o,\ \beta+\underline{y}_o^t G'\underline{x}_o]) =$$
$$= |\alpha\underline{y}_o^t G'\underline{x}_o-\beta\underline{y}_o^t F'\underline{x}_o| \Big/ \sqrt{|\alpha|^2+|\beta|^2}\cdot\sqrt{|\alpha+\underline{y}_o^t F'\underline{x}_o|^2+|\beta+\underline{y}_o^t G'\underline{x}_o|^2} \tag{8.27}$$

□

By using the previous definitions, remarks and propositions we can state the following proposition.

Proposition (8.8): Let the regular pair (F,G) $\in L^{r}_{n,n}$. Let $s_o=\alpha/\beta$ be a single eigenvalue for the pair (F,G) with $\underline{x}_o$, $\underline{y}_o^t$ normalized right and left eigenvectors correspondingly. Then for sufficiently small $F',G'\in\mathbb{R}^{nxn}$ is an eigenvalue s_o' for the pair (F+F',-(G+G')) such that $d(s_o,s_o')\leqq A\sin\theta+B\cos\theta/\nu-E$, where $\theta=\tan^{-1}|\alpha/\beta|$, $A=|\underline{y}_o^t G'\underline{x}_o|$, $B=|\underline{y}_o^t F'\underline{x}_o|$, $\nu=\sqrt{|\alpha|^2+|\beta|^2}$ and $E=\sqrt{B^2+A^2}$.

Proof

From (8.27) we can take,

$$d(s_o,s_o')\leq|\alpha|\cdot|\underline{y}_o^t G'\underline{x}_o|+|\beta|\cdot|\underline{y}_o^t F'\underline{x}_o| \Big/ \sqrt{|\alpha|^2\ |\beta|^2}\cdot\sqrt{|\alpha+\underline{y}_o^t F'\underline{x}_o|^2+|\beta+\underline{y}_o^t G'\underline{x}_o|^2} \tag{8.28}$$

If we put $\theta=\tan^{-1}|\alpha/\beta|$, $\nu=\sqrt{|\alpha|^2+|\beta|^2}$, $A=|\underline{y}_o^t G'\underline{x}_o|$, $B=|\underline{y}_o^t F'\underline{x}_o|$, and $E=\sqrt{B^2+A^2}$, then we can easily seen that $\sin\theta=|\alpha|/\nu$, $\cos\theta=|\beta|/\nu$ and (8.28) becomes

$$d(s_o,s_o')\leqq\nu\sin\theta A+\nu\cos\theta B\Big/\nu\cdot\sqrt{|\alpha+B|^2+|\beta+A|^2}=A\sin\theta+B\cos\theta\Big/\sqrt{|\alpha+B|\ +|\beta+A|^2} \tag{8.29}$$

By using the inequality,

$$\sqrt{|\alpha+B|^2+|\beta+A|^2}\geqq\sqrt{|\alpha|^2+|\beta|^2}-\sqrt{B^2+A^2}, \qquad (A,B\geqq 0) \tag{8.30}$$

(8.29) gives:

$$d(s_o,s_o')\leqq A\sin\theta+B\cos\theta/\nu-E \tag{8.31}$$

□

Remark (8.7): This last inequality in the framework of numerical analysis, says that $1/\nu$ is a condition number for the eigenvalue s_o. □

Remark (8.8): In the case $\theta=0$, then $\sin\theta=0$ and $\alpha=0$ or equivalentrly $s_o=0$. Thus the matrix G is singular; the disappearance of the term $A\sin\theta$ in (8.31) says that perturbations in F cannot affect the singularity of G. □

8.5 Relationships between Φ-(F,G)-e.d.s., deflating subspaces and perturbation results for entirely right regular pairs (F,G)

In the study of perturbation theory of the generalised eigenvalue-eigenvector problem defined on an entirely regular pair $(F,G)\in L^{r}_{n,n}$ a key notion has emerged, the notion of the deflating subspace introduced by Stewart [Ste. -2]. The aim of this section is to establish the links between the invariant subspaces, introduced in the previous chapters and the notion of deflating subspaces. By doing that, the perturbation results established by Stewart for this case may be transferred to the invariant subspaces which have been defined in chapter (7).

Definition (8.7) [Ste. -2]: Let $(F,G)\in L^{r}_{n,n}$ and V be a d-dimensional subspace of $\mathbb{R}^n$. V is a deflating subspace for (F,G) iff $\dim(FV+GV)\leq \dim V=d$. □

Before we examine the exact relationship of a deflating subspace and the invariant subspaces defined before we state the following result.

Proposition (8.9): Let $(F,G)\in L^{r}_{n,n}$ and $V\subset\mathbb{R}^n$ be a d-dimensional subspace. The restriction pencil $(F,G)/V$ has no c.m.i. in its set I_V of strict equivalence invariants.

Proof

Let $(sF-G)V$, V a basis matrix of V, and let $N_r\{sFV-GV\}\neq\{0\}$. Then, $\exists\underline{v}(s)\in\mathbb{R}^n[s]$, $\underline{v}(s)\neq 0$, such that $(sF-G)V\,\underline{v}(s)=\underline{0}$. Define $\underline{x}(s)=V\underline{v}(s)$;

since V has full rank and $\underline{v}(s) \neq \underline{0}$, then $(sF-G)\,\underline{x}(s) = \underline{0}$ and thus sF-G is singular, Q.E.D. □

From the above result, it is clear that $\mathcal{I}_V$ may contain f.e.d., i.e.d. and r.m.i. This observation will be used next.

<u>Theorem (8.2)</u>: Let $(F,G) \in \mathcal{L}^{R}_{n,n}$ and let $\mathcal{V} \subset \mathbb{R}^n$ be a d-dimensional subspace. $\mathcal{V}$ is a deflating subspace iff $\mathcal{V}$ is a Φ-(F,G)-e.d.s. (i.e. $\mathcal{I}_V$ is characterised by e.d. and possibly zero r.m.i.).

<u>Proof</u>

First note that $\dim(F\mathcal{V}+G\mathcal{V}) = \text{rank}_{\mathbb{R}}[FV \,\vdots\, GV]$, where V is any basis matrix of $\mathcal{V}$. Clearly, if $R \in \mathbb{R}^{n\times n}$, $|R| \neq 0$, then $\text{rank}_{\mathbb{R}}[RFV,RGV] = \text{rank}_{\mathbb{R}}[FV,GV]$. We can always choose a special basis $\tilde{V}$ and an R such that $(RF\tilde{V},RG\tilde{V})$ is in the Kronecker form.

By proposition (8.9), the possible set of invariants are e.d. and r.m.i. The typical blocks in $(RF\tilde{V},RG\tilde{V})$ are:

$$RF\tilde{V} = \underbrace{\left[\begin{array}{c:c:c} \multicolumn{3}{c}{0_t} \\ \hdashline \left[\begin{array}{c} I_\zeta \\ \hline \underline{0}^t \end{array}\right] & \multicolumn{2}{c}{0} \\ \hdashline & I_\tau & \\ \hdashline 0 & & H_q \end{array}\right]}_{d}, \qquad RG\tilde{V} = \underbrace{\left[\begin{array}{c:c:c} \multicolumn{3}{c}{0_t} \\ \hdashline \left[\begin{array}{c} \underline{0}^t \\ \hline I_\zeta \end{array}\right] & \multicolumn{2}{c}{0} \\ \hdashline & J_\tau(\alpha) & \\ \hdashline 0 & & I_q \end{array}\right]}_{d} \tag{8.32}$$

where $\left(\begin{bmatrix} I_\zeta \\ \underline{0}^t \end{bmatrix}, \begin{bmatrix} \underline{0}^t \\ I_\zeta \end{bmatrix}\right)$, $(I_\tau, J_\tau(\alpha))$, (H_q, I_q) are typical blocks associated with a r.m.i. ζ,α f.e.d. $(s-\alpha)^\tau$ and an i.e.d. $\hat{s}^q$.

Because of the block diagonal structure in (8.32) the independent columns in $[RF\tilde{V},RG\tilde{V}]$ may be found block by block. Thus, by inspection:

(i) $\left[\begin{array}{c:c} I_{\zeta} & \underline{0}^{t} \\ \underline{0}^{t} & I_{\zeta} \end{array}\right]$ has $\zeta+1$ independent columns for $\forall\ \zeta \in \mathbb{N}$.

(ii) $\left[\begin{array}{c:c} I_{\tau} & J_{\tau}(\alpha) \end{array}\right]$ has τ independent columns for $\alpha \in \mathbb{C}$, $\tau \in \mathbb{N}$.

(iii) $\left[\begin{array}{c:c} H_{q} & I_{q} \end{array}\right]$ has q independent columns for $\forall\ q \in \mathbb{N}$.

and the rank of $[RF\mathcal{V}, RG\tilde{\mathcal{V}}] = \text{rank}[F\mathcal{V}, G\mathcal{V}]$ is equal to $\dim(F\mathcal{V}+G\mathcal{V})$, where

$$\dim(F\mathcal{V}+F\mathcal{V}) = \sum_{i=1}^{t} (\zeta_i+1) + \sum_{i=1}^{\rho} \tau_i + \sum_{i=1}^{\mu} q_i \tag{8.33}$$

However,

$$d = \dim\mathcal{V} = \sum_{i=1}^{t} \zeta_i + \sum_{i=1}^{\rho} \tau_i + \sum_{i=1}^{\mu} q_i \tag{8.34}$$

By (8.33) and (8.34) we have that

$$\dim(F\mathcal{V}+G\mathcal{V}) \geqq \dim\mathcal{V} = d \tag{8.35}$$

and equality holds if and only if there are no non zero r.m.i. in $\mathcal{I}_{\mathcal{V}}$ (zero r.m.i. do not affect the above inequality).

Now if $\mathcal{V}$ is deflating, then for $\dim(F\mathcal{V}+G\mathcal{V}) = d$ all r.m.i. must be zero and thus $\mathcal{V}$ is a Φ-(F,G)-e.d.s. Conversely, if $\mathcal{V}$ is Φ-(F,G)-e.d.s., then all ζ_i are zero and (8.35) holds with equality, i.e. $\mathcal{V}$ is deflating subspace. □

<u>Remark (8.9)</u>: Let $(F,G) \in \mathcal{L}^{h}_{n,n}$ and $\mathcal{V} \subset \mathbb{R}^n$ be a d-dimensional subspace. Then

$$\dim(F\mathcal{V}+G\mathcal{V}) \geqq \dim\mathcal{V} = d \tag{8.36}$$

equality holds true iff $\mathcal{V}$ is a Φ-(F,G)-e.d.s. □

The above remark demonstrates that in the definition of deflating subspaces given by Stewart the equality sign should be used instead of "≦", since there is no subspace of $\mathbb{R}^n$ for which strict inequality of the "<" holds true.

Given that deflating subspaces of an e.r. pair (F,G) may be also expressed for Φ-(F,G)-e.d.s. Thus, following the results of Stewarts [Ste. -2] we may state:

<u>Proposition (8.10)</u>: Let $\mathcal{V}$ be a Φ-(F,G)-e.d.s. Then there are orthogonal matrices K and L such that the first d columns of L span $\mathcal{V}$ and

$$K^tGL = \begin{bmatrix} G_{11} & G_{12} \\ 0 & G_{22} \end{bmatrix}, \quad K^tFL = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix}$$

where G_{11} and F_{11} are dxd matrices.

<u>Proof</u>

Let $L = [L_1 \vdots L_2]$, where $L_1 \in \mathbb{R}^{nxd}$ with orthonormal columns such that colspan $L_1 = \mathcal{V}$ and $L_2 \in \mathbb{R}^{nx(n-d)}$ such tat $L = [L_1 \vdots L_2]$ be orthogonal. Also, let $K_2 \in \mathbb{R}^{nx(n-d)}$ have orthonormal columns lying in the orthogonal complement of $F\mathcal{V}+G\mathcal{V}$ $((F\mathcal{V}+G\mathcal{V})^{\perp})$ and $K_1 \in \mathbb{R}^{nxd}$ be choosen such that $K = [K_1 \vdots K_2]$ is orthogonal.

Since colspan $L_1 = \mathcal{V} \Rightarrow \text{colspan}(GL_1) \subseteq G\mathcal{V} \subseteq F\mathcal{V}+G\mathcal{V}$ and because K_2 is lying in the orthogonal complement of $F\mathcal{V}+G\mathcal{V}$ we have that $K_2^t(GL_1) = 0$. Following the same arguments we have that $K_2^t(FL_1) = 0$.

So,

$$K^tGL = \begin{bmatrix} K_1^t \\ \hline K_2^t \end{bmatrix} G[L_1 \vdots L_2] = \begin{bmatrix} K_1^tGL_1 & K_1^tGL_2 \\ K_2^tGL_1 & K_2^tGL_2 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ 0 & G_{22} \end{bmatrix} \quad \text{and}$$

$$K^tFL = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix}, \text{ thus } K^t(sF-G)L = s\begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix} - \begin{bmatrix} G_{11} & G_{12} \\ 0 & G_{22} \end{bmatrix}$$

□

Also if $\lambda \in \sigma(F_{11}, G_{11})$ with corresponding eigenvector $\underline{z}$, then $\lambda \in \sigma(F,G)$ with corresponding eigenvector $L_1 \underline{z} \in \mathcal{V}$.

□

Next we shall define an operator which plays an important role in deriving the error bounds in the perturbation analysis discussed next.

Definition (8.8)([Ste. -2], [Kato. -1]): Let $G_1, F_1 \in \mathbb{R}^{\ell x \ell}$ and $G_2, F_2 \in \mathbb{R}^{mxm}$ are fixed matrices. For any $X = [P \vdots Q] \in \mathbb{R}^{mx2\ell}$, $P, Q \in \mathbb{R}^{mx\ell}$ we define the operator T on $\mathbb{R}^{mx2\ell}$ as follows:

$$T: \mathbb{R}^{mx2\ell} \rightarrow \mathbb{R}^{mx2\ell}: \forall X = [P \vdots Q] \in \mathbb{R}^{mx2\ell} \rightarrow T(X) \triangleq [PG_1 - G_2Q \vdots PF_1 - F_2Q] \in \mathbb{R}^{mx2\ell} \quad (8.37)$$

$\square$

Lemma (8.1): The operator T is linear.

Proof

Let $X_1 = [P_1 \vdots Q_1]$, $X_2 = [P_2 \vdots Q_2] \in \mathbb{R}^{mx2\ell}$, $\lambda \in \mathbb{R}$, then

$\lambda X_1 + X_2 = [\lambda P_1 + P_2 \vdots \lambda Q_1 + Q_2]$ and $T(\lambda X_1 + X_2) =$

$= [(\lambda P_1 + P_2)G_1 - G_2(\lambda Q_1 + Q_2) \vdots (\lambda P_1 + P_2)F_1 - F_2(\lambda Q_1 + Q_2)] =$

$= [\lambda(P_1G_1 - G_2Q_1) + P_2G_1 - G_2Q_2 \vdots \lambda(P_1F_1 - F_2Q_1) + P_2F_1 - F_2Q_2] =$

$= \lambda[P_1G_1 - G_2Q_1 \vdots P_1F_1 - F_2Q_1] + [P_2G_1 - G_2Q_2 \vdots P_2F_1 - F_2Q_2] = \lambda T(X_1) + T(X_2)$, so T is linear.

$\square$

Lemma (8.2) [Ste. -2]: The operator T is non singular if and only if:

$$\sigma(F_1, G_1) \cap \sigma(F_2, G_2) = \emptyset \quad (8.38)$$

(The spectrum above is defined on the associated pencils $sF_1 - G_1$, $sF_2 - G_2$ in the usual maner).

Remark (8.10): T is non singular is equivalent that,

$\forall [R \vdots S] = Y \in \mathbb{R}^{mx2\ell} \ \exists \ X = [P \vdots Q] \in \mathbb{R}^{mx2\ell}$ such that:

$$T(X) = Y \iff (PG_1 - G_2Q = R \text{ and } PF_1 - F_2Q = S) \quad (8.39)$$

$\square$

Later on we shall need estimates of the size of the solutions of (8.40). In order to do this we define two norms on $\mathbb{R}^{mx2\ell}$ as follows:

Definition (8.9) [Ste. -2]: $\forall\, R = [P \,\vdots\, Q] \in \mathbb{R}^{m\mathrm{x}2\ell}$, where $P,Q \in \mathbb{R}^{m\mathrm{x}\ell}$ we define

$$\|R\|'_2 = \max\{\|P\|_2, \|Q\|_2\} \text{ and } \|R\|'_F = \max\{\|P\|_F, \|Q\|_F\} \tag{8.40}$$

□

We can easily seen that $\|R\|'_2, \|R\|'_F$ are norms on $\mathbb{R}^{m\mathrm{x}2\ell}$.

As we have seen before the operator T is determined by the fixed matrices F_1, G_1, F_2, G_2. Let

$$\mathrm{dif}(F_1,G_1;F_2,G_2) \triangleq \inf_{\|X\|'_F=1} T(X)\|'_F \tag{8.41}$$

Then $\mathrm{dif}(F_1,G_1;F_2,G_2) = 0$ iff T non singular, that is, if and only if $\sigma(F_1,G_1) \cap \sigma(F_2,G_2) = \emptyset$ [Ste. -2].

Also if T is non singular and $T(X) = Y$, then

$$\|X\|'_F \leq \|Y\|'_F \big/ \mathrm{dif}(F_1,G_1;F_2,G_2) \tag{8.42}$$

□

The topic examined next is the study of perturbation properties of Φ-(F,G)-e.d.s., when there is uncertainty in the parameters of the pair (F,G). The following analysis is based on the work of Stewart for delfating subspaces.

Let $L = [L_1 \,\vdots\, L_2]$ and $K = [K_1 \,\vdots\, K_2]$ be orthogonal matrices with $L_1, K_1 \in \mathbb{R}^{n\mathrm{x}\ell}$. Then if $F_{21} = K_2^t F L_1 = 0$ and $G_{21} = K_2^t G L_1 = 0$, the columns of L_1 span an Φ-(F,G)-e.d.s. for sF-G. If F_{21} and G_{21} are small but not exactly zero, it is reasonable to ask whether there exist orthogonal matrices $K' = [K'_1 \,\vdots\, K'_2]$ and $L' = [L'_1 \,\vdots\, L'_2]$ near K and L respectively, such that $K'^t_2 F L'_1 = K'^t_2 G L'_1 = 0$.

We select $K' = [K'_1 \,\vdots\, K'_2] \in S([K_1 \,\vdots\, K_2])$ and $L' = [L'_1 \,\vdots\, L'_2] \in S([L_1 \,\vdots\, L_2])$ in the form

$$\begin{aligned} K'_1 &= (K_1 + K_2 P)(I + P^t P)^{-1/2}, \quad K'_2 = (K_2 - K_1 P^t)(I + PP^t)^{-1/2} \\ L'_1 &= (L_1 + L_2 Q)(I + Q^t Q)^{-1/2}, \quad L'_2 = (L_2 - L_1 Q^t)(I + QQ^t)^{-1/2} \end{aligned} \tag{8.43}$$

where $P, Q \in \mathbb{R}^{(n-\ell)\mathrm{x}\ell}$. It is easy shown that K', L' are orthogonal.

By setting $G_{ij} = K_i^t G L_j$, $F_{ij} = K_i^t F L_j$, $\forall\, i,j = 1,2$, then conditions $K'^t_2 F L'_1 = K'^t_2 G L'_1 = 0$ leads to the following system.

$$\{PG_{11}-G_{22}Q = G_{21}-PG_{12}Q, \quad PF_{11}-F_{22}Q = F_{21}-PF_{21}Q\} \tag{8.44}$$

If we define as before $T([P \vdots Q]) = [PG_{11}-G_{22}Q \vdots PF_{11}-F_{22}Q]$ the system (8.44) becomes:

$$T([P \vdots Q]) = [G_{21}-PG_{12}Q \vdots F_{21}-PF_{12}Q] \tag{8.45}$$

Thus the problem of perturbing K,L into the (deflating) matrices (K',L') is reduced to the following equivalent problem, determine under what conditions the non linear equation (8.45) has a solution. Such conditions are given next.

Proposition (8.11) [Stew. -2]: Let $K = [K_1 \vdots K_2]$, $L = [L_1 \vdots L_2]$ be orthogonal matrices with $K_1, L_1 \in \mathbb{R}^{n x \ell}$. Let also G_{ij}, F_{ij} be defined as before. Let us also define:

$$\gamma = \|[G_{21} \vdots F_{21}]\|_F', \quad \eta = \|[G_{12}^t \vdots F_{12}^t]\|_2', \quad \delta = \mathrm{dif}(F_{11}, G_{11}; F_{22}, G_{22}) \tag{8.46}$$

If $k_1 \triangleq \gamma/\phi^2 < 1/4$, then there are orthogonal matrices $P, Q \in \mathbb{R}^{(n-\ell) x \ell}$ satisfying $\|[P \vdots Q]\|_F' \leqq \gamma/\delta \cdot (1+k) = \gamma/\delta \cdot (1+\sqrt{1-4k_1} \big/ 1-2k_1+\sqrt{1-4k_1}) < 2\gamma/\delta$, such that colspan$(L_1+L_2Q)$ is a deflating and thus a Φ-(F,G)-e.d.s. for the regular pencil sF-G. □

Proposition (8.12): If $F_1, G_1, E_1, E_1' \in \mathbb{R}^{\ell x \ell}$, $F_2, G_2, E_2, E_2' \in \mathbb{R}^{mxm}$, then $\mathrm{dif}(G_1+E_1, F_1+E_1'; G_2+E_2, F_2+E_2') \geq \mathrm{dif}(G_1, F_1; G_2, F_2) - \max\{\|E_1\|_2 + \|E_2\|_2 ; \|E_1'\|_2 + \|E_2'\|_2\}$.

Proof

By definition, $\mathrm{dif}(G_1+E_1, F_1+E_1'; G_2+E_2, F_2+E_2') =$

$$= \inf_{\|[P \vdots Q]\|_F' = 1} \|[P(G_1+E_1)-(G_2+E_2)Q, P(F_1+E_1')-(F_2+E_2')Q]\|_F'$$

Let the infimum be obtained for $X = [P \vdots Q]$, then $\|P\|_F, \|Q\|_F \leqq 1$.

Hence $\mathrm{dif}(G_1+E_1,F_1+E_1';G_2+E_2,F_2+E_2') = \|[P(G_1+E_1)-(G_2+E_2)Q,P(F_1+E_1')-(F_2+E_2')Q]\|_F' =$

$= \max\{\|P(G_1+E_1)-(G_2+E_2)Q\|_F,\|P(F_1+E_1')-(F_2+E_2')Q\|_F\} \geqq$

$\geqq \max\{\|PG_1-G_2Q\|_F-\|E_1\|_2-\|E_2\|_2,\|PF_1-F_2Q\|_F-\|E_1'\|_2-\|E_2'\|_2\} \geqq$

$\geqq \max\{\|PG_1-G_2Q\|_F,\|PF_1-F_2Q\|_F\} - \max\{\|E_1\|_2+\|E_2\|_2,\|E_1'\|_2+\|E_2'\|_2\} \geqq$

$\geqq \mathrm{dif}(G_1,F_1;G_2,F_2) - \max\{\|E_1\|_2+\|E_2\|_2;\|E_1'\|_2+\|E_2'\|_2\}$. $\square$

The question that arises is to investigate whether there exists a V' Φ-(F',G')-e.d.s. of the perturbed pair $(F',G') = (F+E_1,G+E_1')$ which is close to V. Of course the essential question is how close such a subspace may be found with respect to the given perturbation (E_1,E_1'). The answer to this is given next.

<u>Theorem (8.3)</u>: Let $K = [K_1 \vdots K_2]$, $L = [L_1 \vdots L_2]$ be orthogonal matrices with $K_1,L_1 \in \mathbb{R}^{n\mathrm{x}\ell}$, $G_{ij} = K_i^t G L_j$, $F_{ij} = K_i^t F L_j$ $\forall$ $i,j = 1,2$ and suppose that $G_{21} = F_{21} = 0$. (in that case $V = \mathrm{colspan}\ L_1$ is an Φ-(F,G)-e.d.s. of the pair (F,G)). Assume that $E_1,E_1' \in \mathbb{R}^{n\mathrm{x}n}$ and $E_{ij}' = K_i^t E_1' L_j$, $E_{ij} = K_i^t E_1 L_j$ $\forall$ $i,j = 1,2$,

$\varepsilon_{ij} = \max\{\|E_{ij}'\|_F,\|E_{ij}\|_F\}$.

Let us also assume that $\gamma = \varepsilon_{21}$, $\eta = \|[G_{12}^t \vdots F_{12}^t]\|_2' + \varepsilon_{12}$, $\delta = \mathrm{dif}(G_{11},F_{11};G_{22},F_{22}) - \varepsilon_{11} - \varepsilon_{12}$. If $\gamma\eta/\delta^2 < 1/4$ then, there are matrices $P,Q \in \mathbb{R}^{(n-\ell)\mathrm{x}\ell}$ satysfying $\|[P \vdots Q]\|_F'\ 2\gamma/\delta$ such that colspan (L_1+L_2Q) is a Φ-(F,G)-e.d.s. for $s(F+E_1)-(G+E_1')$.

<u>Proof</u>

We will give the proof by using the previous proposition to the problem stated on $s(F+E_1)-(G+E_1')$. If we now denote by $\tilde{G}_{ij},\tilde{F}_{ij}$ it has been denoted by G_{ij}, F_{ij} in the proposition (8.11), then

$$\left.\begin{aligned} \tilde{G}_{ij} &= K_i^t(G+E_1')L_j = K_i^t G L_j + K_i^t E_1' L_j = G_{ij} + E_{ij}' \\ \tilde{F}_{ij} &= K_i^t(F+E_1)L_j = K_i^t F L_j + K_i^t E_1 L_j = F_{ij} + E_{ij} \end{aligned}\right] \quad i,j = 1,2 \qquad (8.47)$$

So $\tilde{G}_{21} = G_{21}+E'_{21} = E'_{21}$, $\tilde{F}_{21} = F_{21}+E_{21} = E_{21}$ and $\tilde{\gamma} = \|[\tilde{G}_{21} \,\vdots\, \tilde{F}_{21}]\|'_F =$

$= \|[E'_{21} \,\vdots\, E_{21}]\|'_F = \max\{\|E'_{21}\|_F, \|E_{21}\|_F\} = \varepsilon_{21}$. Also $\tilde{\eta} = \|[\tilde{G}^t_{12} \,\vdots\, \tilde{F}^t_{12}]\|'_2 =$

$= \max\{\|\tilde{G}^t_{12}\|_2, \|\tilde{F}^t_{12}\|_2\} = \max\{\|G^t_{12}+E'^t_{12}\|_2, \|F^t_{12}+E^t_{12}\|_2\} \leq$

$\leq \max\{\|G^t_{12}\|_2+\|E'^t_{12}\|_2, \|F^t_{12}\|_2+\|E^t_{12}\|_2\} \leqq \max\{\|G^t_{12}\|_2+\|E'_{12}\|_F, \|F^t_{12}\|_2+\|E_{12}\|_F\} \leq$

$\leq \max\{\|G^t_{12}\|_2, \|F^t_{12}\|_2\} + \max\{\|E'_{12}\|_F, \|E_{12}\|_F\} = \|[G^t_{12} \,\vdots\, F^t_{12}]\|'_2 + \varepsilon_{12}$

$\tilde{\delta} = \mathrm{dif}(\tilde{F}_{11}, \tilde{G}_{11}; \tilde{F}_{22}, \tilde{G}_{22}) = \mathrm{dif}(F_{11}+E_{11}, G_{11}+E'_{11}; F_{22}+E_{22}, G_{22}+E'_{22}) \geqq$

$\geqq \mathrm{dif}(G_{11}, F_{11}; G_{22}, F_{22}) - \max\{\|E_{11}\|_2+\|E_{22}\|_2; \|E'_{11}\|_2+\|E'_{22}\|_2\} \geqq$

$\geqq \mathrm{dif}(G_{11}, F_{11}; G_{22}, F_{22}) - \max\{\|E_{11}\|_2, \|E'_{11}\|_2\} - \max\{\|E_{22}\|_2, \|E'_{22}\|_2\} =$

$= \mathrm{dif}(G_{11}, F_{11}; G_{22}, F_{22}) - \varepsilon_{11} - \varepsilon_{12}$.

If we set now:

$\gamma = \varepsilon_{21}$, $\eta = \|[G^t_{12} \,\vdots\, F^t_{12}]\|'_2 + \varepsilon_{12}$, $\delta = \mathrm{dif}(G_{11}, F_{11}; G_{22}, F_{22}) - \varepsilon_{11} - \varepsilon_{12}$, then we have $(\tilde{\gamma} = \gamma,\ \tilde{\eta} \leq \eta,\ \tilde{\delta} \geq \delta) \Rightarrow \tilde{\gamma}\cdot\tilde{\eta}/\tilde{\delta}^2 < \gamma\cdot\eta/\delta^2$, so if we suppose now that $\gamma\cdot\eta/\delta^2 < 1/4$ then $\tilde{\gamma}\cdot\tilde{\eta}/\tilde{\delta}^2 < 1/4$ and the theorem (8.3) is valid for the pencil $s(F+E_1)-(G+E'_1)$; that means that there are matrices $P, Q \in \mathbb{R}^{(n-\ell)\times\ell}$ such that the subspace $\mathcal{V}' = \mathrm{colspan}\ (L_1+L_2Q)$ is a Φ-(F,G)-e.d.s. for $s(F+E_1)-(G+E'_1)$, which of course is a subspace close to the Φ-(F,G)-e.d.s. $\mathcal{V} = \mathrm{colspan}\ L_1$. The closeness can be measured by the gap $(\mathcal{V}, \mathcal{V}') =$ gap (colspan L_1, colspan (L_1+L_2Q)). □

A hint about the possibility of establishing the above result in the case of deflating subspaces was given in [Ste. -2], but no proof was given. The notion of deflating subspaces has been recently extended for general pairs (F,G) by [Van Dor. -3]. The link of this generalised notion of deflating subspaces to the invariant subspaces notions defined in Chapter (7) is still question and it is left to future research.

8.6 The space $L_{m,n}$ as topological vector space (T.V.S.)

The formal analysis of the perturbation properties of matrix pencils, implies the need for the definition of topologies on the set of ordered pairs (F,G). Of course the main question when searching for topologies is to examine which of them is the most suitable for the study of the particular property of the pair (F,G) we would like to examine. Another important factor in determining the suitability of a topology is whether it is related to a natural way to the modelling of the "uncertainty" in the set which is under study. This section serves as an introduction to the study of properties of pairs (F,G) under uncertainty in the parameters of (F,G). We introduce two metric topologies, which are related in a rather natural manner to the modelling of uncertainty on (F,G); in fact it is shown, that the set of pairs $L_{m,n}$ under these two metrics becomes a topological vector space (T.V.S.).

Let $L_{m,n} = \{L\colon L = (F,G),\ F,G \in \mathbb{R}^{m\times n}\}$. Under the standard operations of addition of pairs i.e. $L_1+L_2 = (F_1,G_1)+(F_2,G_2) \triangleq (F_1+F_2,G_1+G_2)$ and scalar multiplication i.e. $\lambda L = \lambda(F,G) = (\lambda F,\lambda G)$, $\lambda \in \mathbb{R}$, $L_{m,n}$ becomes an $\mathbb{R}$-vector space. It is readily shown that $L_{m,n}$ under those operations is a finite dimensional vector space.

Definition (8.10): On the set $L_{m,n}$ we define the functions:

(i) $d\colon L_{m,n}\times L_{m,n} \longrightarrow \mathbb{R}_o^+\colon\ ((F,G),(F',G')) \longrightarrow d((F,G),(F',G')) \triangleq$
$\triangleq \|F-F'\|+\|G-G'\|$

(ii) $d^*\colon L_{m,n}\times L_{m,n} \longrightarrow \mathbb{R}_o^+\colon\ ((F,G),(F',G')) \longrightarrow d^*((F,G),(F',G')) \triangleq$
$\triangleq \max\{\|F-F'\|,\|G-G'\|\}$

where $\|\cdot\|$ is any matrix norm. □

Proposition (8.13): The functions d,d* defined on $L_{m,n}$ are metrics.

Proof

We shall prove the result for d, whereas the proof for d* is similar. In fact, d is a metric on $\mathcal{L}_{m,n}$ because:

(i) $\forall\ (F,G),(F',G') \in \mathcal{L}_{m,n}$ obviously $d((F,G),(F',G')) \geq 0$ and

$$d((F,G),(F',G'))=0 \iff \|F-F'\|+\|G-G'\|=0 \iff F=F' \text{ and } G=G' \iff$$
$$\iff (F,G)=(F',G')$$

(ii) $\forall\ (F,G),(F',G') \in \mathcal{L}_{m,n}$ we have $d((F,G),(F',G'))=\|F-F'\|+\|G-G'\|=$

$$=\|F'-F\|+\|G'-G\|=d((F',G'),(F,G)).$$

(iii) $\forall\ (F,G),(F',G'),(F'',G'') \in \mathcal{L}_{m,n}$ we have that $d((F,G),(F',G'))=$

$$=\|F-F'\|+\|G-G'\|=\|(F-F'')-(F'-F'')\|+\|(G-G'')-(G'-G'')\| \leq$$
$$\leq \|F-F''\|+\|F'-F''\|+\|G-G''\|+\|G'-G''\|=d((F,G),(F'',G''))+d((F',G'),(F'',G'')).$$

So $(\mathcal{L}_{m,n};d)$ and $(\mathcal{L}_{m,n};d^*)$ are metric spaces. $\square$

By using these metrics, the neighborhoods $S_d((F,G),\varepsilon)$, $S_{d^*}((F,G),\varepsilon)$ $\forall\ (F,G) \in \mathcal{L}_{m,n}$ defined by $S_d((F,G),\varepsilon) \triangleq \{(F',G') \in \mathcal{L}_{m,n} : d((F',G'),(F,G))<\varepsilon\}$, $S_{d^*}((F,G),\varepsilon) \triangleq \{(F',G') \in \mathcal{L}_{m,n} : d^*((F',G'),(F,G))<\varepsilon\}$ form a basis for the metric topology on the set $\mathcal{L}_{m,n}$.

Definition (8.11): A sequence $L_n=(F_n,G_n) \in (\mathcal{L}_{m,n};d)$, $n \in \mathbb{N}$ is said to converge in $(\mathcal{L}_{m,n};d)$ iff there exist $L=(F,G) \in (\mathcal{L}_{m,n};d)$ such that, $\forall\ \varepsilon>0\ \exists\, n_o \in \mathbb{N}$: $\forall\ n \geq n_o \Rightarrow (F_n,G_n) \in S_d((F,G),\varepsilon)$ (or equivalently: $d(L_n,L) \to 0$ as $n\to\infty$). We then write $\lim\limits_{n\to\infty}(F_n,G_n)=(F,G)$. $\square$

Remark (8.11): It is readily shown that $\lim\limits_{n\to\infty}(F_n,G_n)=(F,G)$ iff $\lim\limits_{n\to\infty}F_n=F$ and $\lim\limits_{n\to\infty}G_n=G$. $\square$

Proposition (8.14): If $(F_n,G_n) \to (F,G)$ and $(F'_n,G'_n) \to (F',G')$, then in the metric space $(\mathcal{L}_{m,n};d)$, $d((F_n,G_n),(F'_n,G'_n)) \to d((F,G),(F',G'))$.

Proof

Let us assume that $L_n = (F_n, G_n)$, $M_n = (F'_n, G'_n)$, $L = (F,G)$, $M = (F',G')$, $n \in \mathbb{N}$. So we have to prove that: If $L_n \to L$, $M_n \to M$ then $d(L_n, M_n) \to d(L,M)$.

For every $x,y,z,u \in L_{m,n}$ by using the triangle inequality we can see that $|d(x,y) - d(z,u)| \leqq d(x,z) + d(y,u)$.

Thus by setting $x = L_n$, $y = M_n$, $z = L$, $u = M$ we have $|d(L_n, M_n) - d(L,M)| \leqq d(L_n, L) + d(M_n, M)$. From the assumptions $L_n \to L$ and $M_n \to M$ it follows that $d(L_n, L) \to 0$ and $d(M_n, M) \to 0$ and thus $d(L_n, M_n) - d(L,M) \to 0 \Longleftrightarrow$ $\Longleftrightarrow d(L_n, M_n) \to d(L,M)$. □

Remark (8.12): We can derive the same results by using the metric d* on $L_{m,n}$ instead of d. □

Remark (8.13): From the obvious fact that $\max\{\|F-F'\|, \|G-G'\|\} \leqq \|F-F'\| + \|G-G'\|$ we have that $d^*((F,G),(F',G')) \leqq d((F,G),(F',G'))$ $\forall$ $(F,G),(F',G') \in L_{m,n}$; that means that the metric topology induced by d* on $L_{m,n}$ is stronger than the metric topology induced by d. □

Remark (8.14): We can easily seen that the functions,

$+: L_{m,n} \times L_{m,n} \longrightarrow L_{m,n}$: $((F_1,G_1),(F_2,G_2)) \xrightarrow{+} (F_1+F_2, G_1+G_2)$

$\cdot: \mathbb{R} \times L_{m,n} \longrightarrow L_{m,n}$: $(\lambda,(F,G)) \xrightarrow{\cdot} (\lambda F, \lambda G)$

are continuous for the topologies induced by d,d* respectively on $L_{m,n}$. Thus, $(L_{m,n};d)$, $(L_{m,n};d^*)$ are topological vector spaces (T.V.S.).

8.7 The "space-frequency relativistic" properties of $S(F,G)$ and duality

In Chapter (6) the notion of Bilinear strict equivalence E_{H-B} on the set $L_{m,n}$ has been introduced and a complete set of invariants of homogeneous matrix pencils under this equivalence has been defined. The motivation behind this study has been our desire to provide the theory that can highlight the further aspects of duality existing between the differential systems $\mathcal{S}(F,G)$ and $\hat{\mathcal{S}}(F,G)$ which are related in terms of a rather simple bilinear transformation [Kar. & Hay. -1]. In this final section, our effort is focused on the classification of subspaces of the domain of (F,G), as well as of the dynamic properties of $\mathcal{S}(F,G)$ in terms of their invariance, or dependence on $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ transformations. Such a classification may be referred to as a "space-frequency relativistic" classification, since $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$ transformations express coordinate transformations in the domain and codomain of (F,G) and coordinate transformations of the points of the Riemann sphere. The results presented here provide the means for defining "convenient" dual formulations of problems of linear system theory; the term convenient may be referred to the computational, or the conceptual aspects of the considered problem. The duality aspects between the $\mathcal{S}(F,G)$ and $\hat{\mathcal{S}}(F,G)$ differential systems have been already extensively discussed in the previous chapters; the general case of bilinear transformations is examined next.

Definition (8.12): Let $L = (F,G)$, $L' = (F',G') \in \mathcal{L}_{m,n}$ and $\mathcal{S}(F,G): F\dot{\underline{x}} = G\underline{x}$, $\mathcal{S}(F',G'): F\dot{\underline{x}}' = G\underline{x}'$ be the corresponding systems. $\mathcal{S}(F,G)$, $\mathcal{S}(F',G')$ will be called $\mathcal{E}_{\mathcal{H}-\mathcal{B}}$-dual, $\mathcal{E}_{\mathcal{H}}$-dual, if $L\mathcal{E}_{\mathcal{H}-\mathcal{B}}L'$, $L\mathcal{E}_{\mathcal{H}}L'$ respectively. □

All pairs in $\mathcal{L}_{m,n}$ have as a common domain the space $\mathbb{R}^n$. The problem considered next is the study of the nature of a given subspace $\mathcal{V} \subset \mathbb{R}^n$ with respect to different pairs (F,G) of a bilinear orbit $\mathcal{E}_{\mathcal{B}}(F,G)$. By $L_{\mathcal{V}} \overset{\Delta}{=} (F,G)/\mathcal{V}$, $L = (F,G) \in \mathcal{L}_{m,n}$ we shall denote in short the homogeneous restriction pencil. Furthermore, by $\mathcal{H}_{m,n}$ we shall denote the strict equivalence group defined on $\mathcal{L}_{m,n}$

Proposition (8.15): Let $L = (F,G)$, $L' = (F'G') \in \mathcal{L}_{m,n}$, $\mathcal{V} \subset \mathbb{R}^n$ be a subspace and let $\dim\mathcal{V} = d$. If for some $b \in \mathcal{B}$, $L' = b \circ L$, then there exist $h \in \mathcal{H}_{m,d}$ such that $L'_{\mathcal{V}} = h * b \circ L_{\mathcal{V}}$.

Proof

If $L \overset{b}{\to} L' = (F',G')$, where $F' = aFV-cGV$, $G' = dGV-bFV$, then $L_{\mathcal{V}} \overset{b}{\to} L'_{\mathcal{V}} =$ $= \lambda(aFV-cGV)-\mu(dGV-bFV) = \lambda F'V-\mu G'V = L'_{\mathcal{V}}$. Given that V is not uniquely defined, the result follows. $\square$

Remark (8.15): Proposition (8.15) also holds true, if we consider the more general case of $E_{H_{\ell}}-\mathcal{B}$ equivalence where H_{ℓ} is the left strict equivalence subgroup of $H_{m,n}$ (i.e. transformations defined by pair (R,I_n)). $\square$

For a given subspace $\mathcal{V}$ of the domain of $\mathcal{L}_{m,n}$ in general there is no relationship between the strict equivalence invariants of $L_{\mathcal{V}}, L'_{\mathcal{V}}$ for general $L,L' \in \mathcal{L}_{m,n}$. However, if $LE_{\mathcal{B}}L'$, or more generally $LE_{H_{\ell}}-\mathcal{B}L'$, the last result shows that $L_{\mathcal{V}}E_{H-\mathcal{B}}L'_{\mathcal{V}}$ and thus certain relationships hold true between the E_H-invariants of $L_{\mathcal{V}}, L'_{\mathcal{V}}$. These relationships are defined by the properties of E_H-invariants under $E_{\mathcal{B}}$-transformations and according to the nature of $L_{\mathcal{V}}$, the geometric and dynamic properties of $\mathcal{V}$ with respect to different pairs $L \in E_{\mathcal{B}}(F,G)$, the $E_{\mathcal{B}}$-orbit of a given pair, may remain invariant, or vary. The real, complex list and root range of a subspace $\mathcal{V} \subset \mathbb{R}^n$ with respect to a pair $L = (F,G)$, are defined on $L_{\mathcal{V}} = (F,G)/\mathcal{V}$ and shall be denoted by $J_{\mathbb{R}}(L_{\mathcal{V}})$, $J_{\mathbb{C}}(L_{\mathcal{V}})$, $\Phi(L_{\mathcal{V}})$ respectively. In the following the properties of the different types of invariant subspaces under $E_{\mathcal{B}}$-equivalence are examined, identical results may be given for the $E_{H_{\ell}}-\mathcal{B}$-equivalence case.

Proposition (8.16): Let $L = (F,G) \in \mathcal{L}_{m,n}$, $\mathcal{V} \in \mathbb{R}^n$ be a subspace and let $c_{\ell}(F,G;\mathcal{V})$ be the corresponding left (F,G) order of $\mathcal{V}$.

(i) $c_{\ell}(F,G;\mathcal{V})$ is invariant for $\forall L' = (F',G';\mathcal{V}) \in E_{\mathcal{B}}(F,G)$.

(ii) If $\mathcal{V}$ is a $(\mathcal{W},\mathcal{U})$-p.i.s. with respect to (F,G) then it is also a $(\mathcal{W},\mathcal{U})$-p.i.s. with respect any $(F',G') \in E_{\mathcal{B}}(F,G)$.

Proof

$c_\ell(F,G;V)$ is devined by the set of r.m.i. and thus it is invariant under E_B-equivalence, which proves (i). The subspace V is a (W,U)-p.i.s., iff $c_\ell(F,G;V)=0$. By part (i) the result follows. $\square$

The property of V to be a (W,U)-p.i.s. is invariant under E_B-equivalence and thus it is a property of the orbit $E_B(F,G)$ and not only of the particular $L \in E_B(F,G)$. The family of all possible (W,U)-p.i.s. defined on (F,G) shall be denoted by $Z_{p.i.}$. The above result also implies that for specific subfamilies of $Z_{p.i.}$, their more specific characterising property always varies within the set of properties characterising subfamilies of $Z_{p.i.}$ under E_B-equivalence. The E_B-invariance property of more specific subfamilies of $Z_{p.i.}$ is considered next.

Theorem (8.4): Let $L=(F,G) \in L_{m,n}$, $V \subset \mathbb{R}^n$ be a c.-(F,G)-i.s. and let $J_{\mathbb{R}}(L_V)$ be the corresponding real list. Necessary and sufficient condition for V to be a c.-(F',G')-i.s. for $\forall\ (F',G') \in E_B(F,G)$, is that $J_{\mathbb{R}}(L_V)=\emptyset$.

Proof

Since V is a c.-(F,G)-i.s., then it may be decomposed as $V=V_\varepsilon \oplus V_\alpha$ where V_α may be expressed as $V_\alpha = V_\alpha^{\mathbb{R}} \oplus V_\alpha^{\mathbb{C}}$; $V_\alpha^{\mathbb{R}}, V_\alpha^{\mathbb{C}}$ are finite non zero e.d.s. which correspond to real, complex e.d. respectively.

If $J_{\mathbb{R}}(L_V)=\emptyset$, then the invariance of c.m.i. implies that V_ε will be a c.m.i. subspace for every $(F',G') \in E_B(F,G)$. Similarly, because a pair of complex conjugate e.d. is always mapped under any $b \in B$ to a pair of complex conjugate e.d., it follows that under any $b \in B$ $V_\alpha^{\mathbb{C}}$ becomes also a $V_{\alpha'}^{\mathbb{C}}$ subspace which is c.-(F,G)-i.s.. Therefore, if $J_{\mathbb{R}}(L_V)=\emptyset$, then V is c.-(F,G)-i.s. for $\forall\ L \in E_B(F,G)$. The necessity is proved by contradiction. Assume that $J_{\mathbb{R}}(L_V) \neq \emptyset$. Then there exist a real non zero e.d.s. in $V_\alpha^{\mathbb{R}}$, say $\mathcal{V}_{\lambda}^{\mathbb{R}}$, associated with a frequency $\lambda \in \mathbb{R}-\{0\}$. We can always define a $b \in \mathcal{B}$ such that λ may be mapped either at zero, or infinity; Under such transformations $\mathcal{V}_{\lambda}^{\mathbb{R}}$ will behave either as a zero e.d.s., or as an infinite e.d.s. Clearly, for some appropriate $b \in \mathcal{B}$ the c.-(F,G)-i.s. property is violated. □

A c.-(F,G)-i.s. $\mathcal{V}$ with $J_{\mathbb{R}}(L_{\mathcal{V}})$ will be referred to as a strong complete-(F,G)-invariant subspace (s.c.-(F,G)-i.s.) and the family of all such subspaces defined on (F,G) will be denoted by $\mathcal{V}_{s.c.}$. The property of a subspace to be a c.-(F,G)-i.s. is clearly invariant for subspaces of the $\mathcal{V}_{s.c.}$ under $\mathcal{E}_{\mathcal{B}}$-equivalence. Two important subfamilies of $\mathcal{V}_{s.c.}$ are: the family $\mathcal{P}_{i.r}$ of all $\mathcal{I}_c$-(F,G)-c.m.i.s. and the family $\mathcal{F}_{f.i.}$ of Φ-(F,G)-e.d.s. with $J_{\mathbb{R}}(L_{\mathcal{V}}) = \emptyset$; a subspace $\mathcal{V} \in \mathcal{F}_{f.i.}$ will be called a strong-Φ-(F,G)-e.d.s. (s.-Φ-(F,G)-e.d.s.). From Theorem (8.4) we have:

Corollary (8.1): Let $L = (F,G) \in \mathcal{L}_{m,n}$ and let $\mathcal{V} \subset \mathbb{R}^n$ be a subspace.

(i) If $\mathcal{V}$ is an $\mathcal{I}_c$-(F,G)-c.i.s., then for $\forall$ $(F',G') \in \mathcal{E}_{\mathcal{B}}(F,G)$ $\mathcal{V}$ is also an $\mathcal{I}_c$-(F',G')-c.i.s.

(ii) If $\mathcal{V}$ is a s.-Φ-(F,G)-e.d.s. with list $J_{\mathbb{C}}(L_{\mathcal{V}})$, then for $\forall$ $(F',G') \in \mathcal{E}_{\mathcal{B}}(F,G)$ $\mathcal{V}$ is also a s.-Φ'-(F,G)-e.d.s. with list $J_{\mathbb{C}}(L_{\mathcal{V}}) = J_{\mathbb{C}}(L'_{\mathcal{V}})$. □

Note that under $\mathcal{E}_{\mathcal{B}}$-equivalence the lists $J_{\mathbb{R}}(L_{\mathcal{V}})$, $J_{\mathbb{C}}(L_{\mathcal{V}})$ are invariant, but not the root range.

Corollary (8.2): Let $(F,G) \in \mathcal{L}_{m,n}$, $\mathcal{E}_{\mathcal{B}}(F,G)$ its $\mathcal{E}_{\mathcal{B}}$-orbit and let $\mathcal{V} \subset \mathbb{R}^n$ be a subspace. If $\mathcal{V}$ is e.-(G,F)-i.s., or e.-(F,G)-i.s., or e.c.-(F,G)-i.s., then $\exists$ $(F',G') \in \mathcal{E}_{\mathcal{B}}(F,G)$ such that $\mathcal{V}$ is a c.-(F',G')-i.s. □

The proof of the above result follows along similar lines to that of Theorem (8.4). From the above two corollaries we may give a classification of the dynamic properties of a subspace $\mathcal{V}$ under $\mathcal{E}_{\mathcal{B}}$-equivalence transformations.

Corollary (8.3): Let $(F,G)\in \mathcal{L}_{m,n}$, $\mathcal{E}_{\mathcal{B}}(F,G)$ its $\mathcal{E}_{\mathcal{B}}$-orbit and let $\mathcal{V}\subset \mathbb{R}^n$ be a subspace.

(i) If $\mathcal{V}$ is a C^{∞}-r.s. with respect to (F,G), then for $\forall\ (F',G')\in \mathcal{E}_{\mathcal{B}}(F,G)$ $\mathcal{V}$ is a C^{∞}-r.s.

(ii) If $\mathcal{V}$ is a $D'_{\mathcal{B}}$-h.s. with respect to (F,G), then for $\forall\ (F',G')\in \mathcal{E}_{\mathcal{B}}(F,G)$ $\mathcal{V}$ is a $D'_{\mathcal{B}}$-h.s.. There exist, however, $(F',G')\mathcal{E}_{\mathcal{B}}(F,G)$ such that $\mathcal{V}$ is C^{∞}-h.s. with respect to (F'',G'').

(iii) If $\mathcal{V}$ is a $D'_{\mathcal{B}}$-r.s. with respect to (F,G), there exist $(F',G')\in \mathcal{E}_{\mathcal{B}}(F,G)$ such that $\mathcal{V}$ is not a $D'_{\mathcal{B}}$-r.s. with respect to (F',G'). $\square$

From the above result it follows that the C^{∞}-reachability and $D'_{\mathcal{B}}$-holdability are "strong" properties since they hold true with respect to any pair in $\mathcal{E}_{\mathcal{B}}(F,G)$ and any general subspace $\mathcal{V}$ having the above properties. Contrary to that, the C^{∞}-holdability and $D'_{\mathcal{B}}$-reachability are "weak" properties since for a general subspace $\mathcal{V}$ either of the above properties may depend on the particular $(F',G')\in \mathcal{E}_{\mathcal{B}}(F,G)$.

8.8 Conclusions

The topological results given in this chapter, are of a preliminary nature. In fact, they connect some of the known results on the perturbation theory of the generalised-eigenvalue eigenvector problem with the geometric concepts which have been developed in the previous section. The means by which these links have been achieved is via the new metrics which have been defined. The importance of these new metrics in the study of properties of pencils under uncertainty has yet to be proved; however, the links with the standard theory, their "easy" from the computational viewpoint nature and their intuitively simple forms, as means to measure uncertainty are encouraging indicators.

The application of the Bilinear Strict equivalence theory for the classification of the geometric and dynamic properties of the various types of invariant subspaces, under $\mathcal{E}_{\mathcal{B}}$-equivalence, provide a "relativistic" classification of these properties. Thus, the notions of C^{∞}-reachability and $D'_{\mathcal{B}}$-holdability have emerged as "strong" properties, whereas those of $D'_{\mathcal{B}}$-reachability and C^{∞}-holdability as "weak". The results on $\mathcal{E}_{\mathcal{B}}$-equivalence may be used for the definition of convenient dual systems in the standard, or extended state space theory. In the standard linear system thoery, the above "relativistic" classification implies that controllability, observability, and almost (A,B)-invariance are "strong" properties, whereas those of almost controllability and (A,B)-invariance are "weak"; furthermore it is worth pointing out that in this context, the property of stability is also weak, since it depends on the particular transformation $b \in \mathcal{B}$.

CHAPTER 9:

Conclusions

CHAPTER 9: CONCLUSION, FUTURE RESEARCH

The main objective of the thesis was to develope further a number of important aspects of matrix pencil theory which are relevance to linear systems thoery. By creating an enriched matrix pencil theory it is believed that a unifying, matrix pencil based approach to the study of regular and extended state space systems may be established. The results in this thesis aim at this direction. In fact, they establish a complete number theoretic treatment of the S.E.invariants, develope further the geometric theory of matrix pencils and the dynamic theory of $S(F,G)$ systems, produce a framework for the study of stability of invariants (robustness) and develope a theory of invariants under B.S.E. transformations.

The essence of the number theoretic characterisation of the S.E. invariants is that it is based on the study of Piecewise Arithmetic Progression type sequences defined on an ordered pair (F,G), without using the underlying algebraic notions. Given that all types of invariants are characterised in a similar manner, the approach is unifying. For the computation of the sequences it is assumed that the root range of (F,G) is known; a singular value decomposition may then be used to compute the ranks of appropriate sequences of matrices, and from those ranks the sequences. The analysis of discontinuity properties may be carried out by Ferrer's type diagrams. These results lead to a method for computing the Kronecher form without using special type transformations. The only inherent computational difficulty of the method is defining the root range of (F,G).

The study of geometry of matrix pencils presented here has two interelated parts. The first deals with the geometry of the different types of strict equivalence invariants and it is manifested by the structure of generalised null spaces (the case of e.d.) an annihilating spaces (case of c.m.i., r.m.i.). This part is intimately connected with the P.A.P.S. theory and defines the structure of bases of the elementary invariant subspaces (i.e. those characterised by one type of S.E. invariants). The properties of the special basis matrices indicate the nature of the abstract subspace algorithms developed in geometric theory; in fact, it is believed, that it is possible to develope these abstract algorithms using the properties of the basis matrices itself. An additional gain out of this study is the understanding of the module structure of the right and left null spaces of a pencil; these result culminate in a purely geometric construction of the minimal bases. It is believed that the minimal bases results, derived on matrix pencils may be extended to the case of a general polynomial matrix, thus providing a geometric theory of minimal bases with obvious computational advantages.

The second part of the geometry study revolves around the restriction pencil of a given subspace. The notions of (F,G)-, (G,F)-, complete (F,G)-invariant subspaces, extended (F,G)-, (G,F)-, complete (F,G)-invariant subspaces and (F,G)-right annihilating spaces generalise the standard notions of invariant subspaces of the geometric theory, since first they extend them to the case of a general pair (F,G) (in the standard geometric theory (F,G) is entirely right singular) and second because they also imply the duals of the standard geometric theory invariant subspaces. The advantage of the new notions is due to the generality of the pair, their close links with the number theoretic and computational aspects of (F,G), as well as with the underlying algebra. The asymptotic properties of infinite spectrum invariant subspaces may be established by using the notion of invariant regular realizations $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ of the triple $(F,G;\mathcal{V})$; this approach enlarges the domain and the sequences of subspaces are from the domain of $(\tilde{F},\tilde{G})$. It is questionable whether an asymptotic characterisation of infinite spectrum subspaces may be achieved with sequences from the domain of (F,G).

The theory of invariant regular realizations of a triple $(F,G;\mathcal{V})$ may be extended to the case where the realization $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ has $(\tilde{F},\tilde{G})$ not entirely right singular; in fact, $(\tilde{F},\tilde{G})$ may be characterised by c.m.i. and i.e.d. Such an invariant theory allows the bridging of the general theory, presented on (F,G), with extended state space linear systems thory. There are however, a number of problems which have to be resolved here; especially those related to the interpretation of S.E. transformations with meaningful notions of feedback.

The dynamic propertis of $\mathcal{S}(F,G)$, i.e. the properties of the solution space, have demonstrated that the S.E. invariants are essential in the characterisation of the redundancy space, the initial space and solution space of the differential system. They have emphasised that system theoretic properties such as controllability, almost controllability have a deeper meaning, since they are also valid for $\mathcal{S}(F,G)$ representations; in fact, for the latter case they take the form of C^{∞}-, distributional reachability. The different families of invariant subspaces have been classified according to the properties of C^{∞}-, distributional holdability and C^{∞}-, distributional reachability; these properties are the counterparts of (A,B)-invariance, almost (A,B)-invariance and controllability, almost controllability. As with the asymptotic characterisation of infinite spectrum invariant subspaces, it is really questionable, whether distributional holdability, reachability may be interpreted as almost C^{∞}-holdability, reachability properties (in the sense defined by Willems in standard geometric theory) with trajectories in the domain of (F,G). Using the regular invariant realization, however, the almost C^{∞}-holdability, reachability property may be established for the subspace $\tilde{\mathcal{V}}$ of the realization $(\tilde{F},\tilde{G};\tilde{\mathcal{V}})$ of the triple $(F,G;\mathcal{V})$. Investigating the asymptotic characterisation of infinite spectrum invariant subspaces and the almost C^{∞}-holdability, reachability properties of distritonally holding, reachability subspaces of (F,G), on (F,G) itself without expanding the domain with an invariant regular realization, is a problem that deserves further investigation.

The importance of the perturbation properties of S.E. invariants and invariant subspaces of a pair (F,G) needs hardly to be emphasized. The link of the metric topologies defined in Chapter (8) with standard results of the perturbation theory of the generalised eigenvalue eigenvector problem is encouraging; the richness of the theory of topological vector spaces provides a fruther encouragement för trying to develope the robustness aspects of the invariants of (F,G) along these lines. The type of results we are interested in, are a general theory of robustness of invariant subspaces (along lines similar to those given for the deflating subspaces of a regular pair) and a systematic study of the generic properties of a pair (F,G). It is believed that the general notion of deflating subspaces of a general pair (F,G) [Van Do. -3] is related to a specific way to the invariant subspaces discussed in this thesis. The exact determination of this characterisation is one of the first priorities in a future research.

The theory of invariants of matrix pencils under B.S.E. transformations provides the means for the "space-frequency" relativistic classification of geometric and dynamic properties of (F,G). Applying the theory to the case of constructing convenient dual problems in linear systems is considered as an important area for future research. The question of defining a canonical form for B.S.E., seems to be a rather hard one; it is a theoretical one and it is connected with the construction of canonical dual problem, i.e. selection of appropriate "space-frequency" setting, which demonstrates the invariant structure of (F,G) and thus of the associated problem. Specialising the various types of B.S.E. invariants to the set of ordered pairs (F,G) (a necessary step in the definition of canonical blocks) is rather difficult, because of the hard nature of B.S.E. invariants. However, it is still an interesting area that deserves further consideration. The work of Kublanovskaya [Kubl. -1] on the conditioning of the generalised eigenvalue-eigenvector problem with simple bilinear transformations suggests that B.S.E. is of immense importance from the numerical analysis viewpoint. Defining an "optimal" bilinear transformation, that will create the best conditioning for computations is of crucial importance and it is one of the problems we consider for future research. The condition number of (F,G) does not belong to the set of B.S.E. invariants and thus its assignment in an optimal way may be possible.

Presenting the theory of S.E. on the pair (F,G) in terms of the number theoretic properties of sequences is a more natural way for extending the theory to cases where the elements of F,G are not from $\mathbb{R}$ or $\mathbb{C}$, but for more general fields, or possibly rings. Such a theory does not exist at the moment. Developing a theory of strict equivalence for more general case pair (F,G) is a prerequisite for expanding the matrix pencil theory to linear time varying systems and singularly perturbed systems. The underlying motivation behind such an approach, is that the ordered pair operator description is a natural representation of first order linear differential equations, thus, time varying and singularly perturbed linear systems may be treated with a generalised matrix pencil theory, if such a theory is available.

REFERENCES

[Apl. -1] : Aplevich, D.J. "Minimal Representations of Implicit Linear Systems". Automatica, Vol. 21, No 3, pp. 259-269, 1985.

[Apl. -2] : Aplevich, D.J. "Time-domain input-output representations of linear systems". Automatica, 17, 509, 1981.

[Bas. & Mar. -1] : Basile, G. and Marro, G. "Controlled and conditioned invariant subspaces in linear system theory". J.Optimiz. Theory Appl., Vol. 3, pp. 306-315, 1969.

[Ber. -1] : Bernhard, P. "On singular implicit linear dynamical systems", SIAM J.Control Optimization, 20, 612, 1982.

[Bir. & MacL. -1] : Birkhoff, G. and MacLane, G. "Algebra", Collier-MacMillan Limited, London, 1967.

[Bru. -1] : Brunovsky, P. "A classification of linear controllable systems", Kybernetica, 3, 173-187, 1970.

[Cal. & Des. -1] : Callier, F.M. and Desoer, C.A. "Multivariable Feedback Systems", Springer-Verlag, New York, 1982.

[Camp. -1] : Campbell, S.L. "Singular systems of differential equations", Research Notes in Mathematics, Vol. I, Pitman Advanced Publishing Program, 1980.

[Camp. -2] : Campbell, S.L. "Singular systems of differential equations", Research Notes in Mathematics, Vol. II, Pitman Advanced Publishing Program, 1981.

[Cob. -1] : Cobb, J.D. "On the solutions of linear differential equations with singular coefficients", Preprint, 1980.

[Doe. -1] : Doetsch, G. "Introduction to the theory and application of the Laplace transformation", Springer-Verlag, 1970.

[Dor. -1] : Dorny, C. "A vector space approach to models and optimization", John Wiley & Sons, 1975.

[Forn. -1] : Forney, D.G. "Minimal bases of rational vector spaces with applications to multivariable linear systems", SIAM J. Control, 13, 493-520, 1975.

[Fran. -1] : Francis, B.A. "Convergence in the bountary layer for singularity perturbed equations", Automatica, 1982.

[Gant. -1] : Gantmacher, F.R. "The thoery of matrices", 2 Vols., Chelsea, New York, 1959.

[Gian. -1] : Giannakopoulos, Ch. "Frequency assignment problems of linear multivariable systems: An exterior algebra and algebraic geometry based approach", Ph.D. Thesis, The City University, London, 1984.

[Gre. -1] : Greub, W.H. "Multilinear Algebra", Springer-Verlag, 1967.

[Gru. & Wei. -1] : Gruenberg, K.W. and Weir, A.J. "Linear Geometry", Springer-Verlag, 2nd edition, New York, 1977.

[Hir. & Smal. -1] : Hirsch, M. and Smale, S. "Differential equations, dynamical systems and linear Algebra", Academic Press, 1974.

[Hod. & Ped. -1] : Hodge, W.V.D. and Pedoe, D. "Methods of Algebraic Geometry", 3 Vols., Cambridge University Press, 1952.

[Jaf. & Kar. -1] : Jaffe, S. and Karcanias, N. "Matrix pencil characterization of almost (A,B)-invariant subspaces: A classification of geometric concepts", Intern.J.Control, Vol. 33, No 1, pp. 51-93, 1981.

[Kag. & Ruh. -1] : Kagström, B. and Ruhe, A. (Edit.): "Matrix pencils". Conf. Proc., Pite Harsbad, Sweden (1982). Published in Lecture Notes in Math., No 973, Springer-Verlag, Berlin, 1983.

[Kail. -1] : Kailath, T. "Linear systems", Prentice Hall, New Jersey, 1980.

[Kal. -1] : Kalman, R.E. "Kronecker invariants and feedback". Proc. of Conf. on ordinary differential equations (Ed. L.Weiss) NLR Math.Res.Cent., 459-471, 1965.

[Kar. -1] : Karcanias, N. "Matrix pencil approach to geometric system theory", Proc. IEE, 126, 585-590, 1979.

[Kar. -2] : Karcanias, N. "On the basis matrix characterisation of controllability subspaces", Intern.J.Control, Vol. 29, pp. 767-786, 1979.

[Kar. -3] : Karcanias, N. "The use of Grassmann products in the computation of the invariant zero polynomial of square linear proper systems", Intern.J.Control, Vol. 28, pp. 361-381, 1978.

[Kar. -4] : Karcanias, N. "Course notes on Exterior Algebra", Control Engineering Centre, The City University, 1983.

[Kar. -5] : Karcanias, N. "Notes on matrix pencil theory", Unpublished Notes, Control Engin.Centre, The City University, 1981.

[Kar. & Hay. -1] : Karcanias, N. and Hayton, G.E. "Generalised autonomous differential systems, algebraic duality and geometric theory", Proc. IFAC VIII, Triennial World Congress, Kyoto, Japan, August, 24-28, 1981.

[Kar. & Hay. -2] : Karcanias, N. and Hayton, G.E. "Generalised autonomous differential systems, algebraic duality and geometric theory", Research Report, The City University, Systems Science, DSS/N.K./G.E.H./212, 1982.

[Kar. & Hay. -3] : Karcanias, N. and Hayton, G.E. "State-space and transfer function invariant infinite zeros: A unified approach. Proc.Joint Aut.Control Conf.Univ.of Virginia, Charlottesville, June, 17-19, 1981.

[Kar. & Kal. -1] : Karcanias, N. and Kalogeropoulos, Gr. "Bilinear-strict equivalence of matrix pencils and autonomous singular differential systems", Published in the Proceed. of 4th I.M.A. Intern.Conf. on Control Theory, Robinson College, Cambridge, 1984.

[Kar. & Kouv. -1] : Karcanias, N. and Kouvaritakis, B. "The output zeroing problem and its relationship to the invariant zero structure", Inter. J.Control, 30, 395-415, 1979.

[Kar. & Kouv. -2] : Karcanias, N. and Kouvaritakis, B. "The use of frequency transmission concepts in linear multivariable system analysis", Inter.J.Control, Vol. 28, No 2, 197-240, 1978.

[Kar. & Kouv. -3] : Karcanias, N. and Kouvaritakis, B. "Zero time adjustement of initial conditions and its relationship to controllability subspaces", Inter.J.Control, Vol. 29, No 5, 749-765, 1979.

[Kar. & MacB. -1] : Karcanias, N. and MacBean, P. "Structural invariants and canonical forms of linear multivariable systems", University of Cambridge, Research Report, CUED/F-CAMS/TR-192, 1979, Third IMA Conference on Control Theory, Acad. Press, 257-282, 1981.

[Kat. -1] : Kato, T. "Perturbation theory for linear operators", Springer-Verlag, New York, 1976.

[Kouv. & Edm. -1] : Kouvaritakis, B. and Edmunds, J. "Multivariable root loci; a unified approach to finite and infinite zeros", Int.J. Control, Vol. 29, 393-428, 1979.

[Kubl. -1] : Kublanovskaya, V.N. "On algorithms for the solution of spectral problems of linear matrix pencils", LOMI preprints E-10-81, Leningrad, 1981.

[Loi. -1] : Loiseau. "Some geometric considerations about the Kronecker normal form", To appear in Inter.J.Control, 1985.

[Lue. -1] : Luenberger, D.C. "Dynamic equations in descriptor form", IEEE Trans.Aut.Control, AC-22, 312-321, 1977.

[MacF. & Kar. -1] : MacFarlane, A.G.J. and Karcanias, N. "Relationships between state space and frequency respons concepts", Proc. 7th IFAC Congr., Part 43, Helsinki, 1978.

[MacF. & Kar. -2] : MacFarlane, A.G.J. and Karcanias, N. "Poles and zeros of linear multivariable systems: a survey of the algebraic, geometric and complex variable theory", Inter.J. Control, Vol. 24, 33-37, 1976.

[Mar. -1] : Marcus, M. "Finite dimensional multilinear algebra" (in two parts). Marcel and Deker, New York, 1973.

[Mar. & Min. -1] : Marcus, M. and Minc, H. "A survey of matrix theory and matrix inequalities", Allyn and Bacon, Boston, 1964.

[Mor. -1] : Morse, A.S. "Structural invariants of lineal multivariable systems", SIAM J.Control, 11, 446-465, 1973.

[Post. & MacF. -1]: Postlethwaite, I. and MacFarlane, A.G.J. "A complex variable approach to the analysis of linear multivariable feedback systems", Lecture Notes in Control and Information Sciences, Springer-Verlag, 1979.

[Ro. -1] : Rosenbrock, H.H. "State-space and multivariable theory", Nelson, 1970.

[Ro. -2] : Rosenbrock, H.H. "Structural properties of linear dynamical systems", Inter.J.Control, 20, 191-202, 1974.

[Ro. & Hay. -1] : Rosenbrock, H.H. and Hayton, G.E. "Dynamical indices of a transfer function matrix", Inter.J.Control, Vol. 20, pp. 177-189, 1974.

[Se. & Kn. -1] : Semple, G.J. and Kneebone, T.G. "Algebraic Projective Geometry", Oxford at the Clarendon Press, 1952.

[Smi. -1] : Smith, M.C. "A generalised Nyquist/root-locus theory for multi-loop feedback systems", Ph.D.thesis, Cambridge University, 1982.

[Ste. -1] : Stewart, G.W. "Error and perturbation bounds for subspaces associated with certain eigenvalue problems", SIAM Rev., 15, 727-769, 1973.

[Ste. -2] : Stewart, G.W. "Perturbation theory for the generalized eigenvalue problem, recent advances in numerical analysis", Ed. C.De Boor, G.Golub, Academic Press, New York, 193-206, 1978.

[Sun. -1] : Sun, J.g. "Perturbation analysis for the generalised eigenvalue and the generalised singular value problem. Matrix pencils", Lecture Notes in Mathematics, Springer-Verlag, 1982.

[Sun. -2] : Sun, J.g. "Invariant subspaces and generalized invariant subspaces", (I),(II) Math.Num.Sinica, 2, 113-123, 1980.

[Thor. -1] : Thorp, J.P. "The singular pencil of linear dynamical systems", Inter.J.Control, 18, 577-596, 1973.

[Trent. -1] : Trentelman, H.L. "Almost invariant subspaces and high gain feedback", Ph.D.Thesis, Department of Mathematics and Computing Science, Eindhoven University of Technology, 1985.

[Tur. -1] : Turnbull, H.W. "The theory of determinants, matrices and invariants", Blackie & Son Ltd., London, 1928.

[Tur. & Ait. -1] : Turnbull, H.W. and Aitken, A.C. "An introduction to the theory of canonical matrices", Dover Publications, 1961.

[VanDo. -1] : VanDooren, P. "The computation of Kronecker's canonical form of a singular pencil", Lin.Alg. & Appl., Vol. 27, 103-141, 1979.

[VanDo. -2] : VanDooren, P., Dewilde, P. and Vandewalle, J. "On the determination of the Smith-McMillan form of a rational matrix from it's Laurent expansion", IEEE Trans.Circuits and Systems, 1979.

[VanDo. -3] : VanDooren, P. "Reducing subspaces: definitions, properties and algorithms", Matrix Pencils, Lecture Notes in Mathematics, No 973, Springer-Verlag, 1983.

[Ver. -1] : Verghese, G "Infinite frequency behaviour in generalised dynamical systems", Ph.D.Thesis, Stanford University USA, 1978.

[War. & Eck. -1] : Warren, M.E. and Eckberg, A.E. "On the dimensions of controllability subspaces: A characterisation via polynomial matrices and Kronecker invariants", SIAM J.Control, 13, 434-445, 1975.

[Wil. -1] : Willems, J.C. "Almost invariant subspaces. An approach to high gain feedback design, part I,II: Almost controlled subspaces", IEEE Trans.on Aut.Control, Vol. AC-26, 235-252, 1981, Bol. AC-27, 1071-1085, 1982.

[Wil. -2] : Willems, J.C. "Almost A (mod B) invariant subspaces", Asterisque, Vol. 75-76, pp. 239-248, 1980.

[Wilk. -1] : Wilkinson, J.H. "The algebraic eigenvalue problem", Clarendon Press, Oxford, 1965.

[Wilk. -2] : Wilkinson, J.H. "Linear differential equations and Kronecker's canonical form. Recent advances in numerical analysis" (C.deBoor and G.Golumb, Eds.), Academic Press, New York, 1978.

[Won. -1] : Wonham, W.M. "Linear multivariable control: a geometric approach", Second Edition, Springer-Verlag, 1979.

[Won. & Mor. -1] : Wonham, W.M. and Morse, A.S. "Decupling and pole assignment in linear multivariable systems: a geometric approach", SIAM J.Control, 8(1), pp. 1-18, 1970.